DOCUMENT RESUME

ED 483 034 UD 035 692

TITLE	Research Perspectives on School Reform: Lessons from the Annenberg Challenge.
INSTITUTION	Brown Univ., Providence, RI. Annenberg Inst. for School Reform.
PUB DATE	2003-03-00
NOTE	136p.
AVAILABLE FROM	Annenberg Institute for School Reform, Brown University, BOx 1985, Providence RI 02912. Web site: http://www.annenberginstitute.org.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	EDRS Price MF01/PC06 Plus Postage.
DESCRIPTORS	*Academic Achievement ; *Accountability; *Educational Policy; Elementary Secondary Education; *Incentives; State Government

ABSTRACT

This collection of essays is the latest contribution to the body of cross-site knowledge on the Annenberg Challenge. The intent of this publication is to further advance the understanding and practice of educators and policy leaders about what it takes to achieve and sustain large-scale reform, all the more urgent in light of new state and federal mandates. (Author)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 483 034

Research Perspectives on School Reform:

LESSONS FROM THE ANNENBERG CHALLENGE

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Susan Fisher

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Annenberg Institute for School Reform at Brown University

UD 033 781

2

BEST COPY AVAILABLE

Research Perspectives on School Reform:

LESSONS FROM THE ANNENBERG CHALLENGE

Annenberg Institute for School Reform at Brown University

Credits

Design:
Gilbert Design Associates

Cover photo:
courtesy of Transforming
Education Through the
Arts Challenge

March 2003

Table of Contents

5 FOREWORD
Warren Simmons & Michael Grady, Annenberg Institute for School Reform

9 INTRODUCTION
Brenda J. Turnbull, Policy Studies Associates

15 EVOLUTION IN TANDEM: DEVELOPMENT AND RESEARCH IN AN ARTS EDUCATION PROGRAM
Hollis Headrick & Greg McCaslin, The Center for Arts Education
Terry Baker, Education Development Center, Inc.

35 BUILDING CONSTRUCTIVE PARTNERSHIPS IN URBAN SCHOOL REFORM
Joy Phillips & Pedro Reyes, University of Texas at Austin
Linda Clarke, Houston Annenberg Challenge

51 TRANSFORMING EVENTS: A LOCAL EDUCATION FUND'S EFFORTS TO PROMOTE LARGE-SCALE URBAN SCHOOL REFORM
Barbara Neufeld, Education Matters, Inc.
Ellen Guiney, Boston Plan for Excellence in the Public Schools

69 BALANCING SUMMATIVE AND FORMATIVE EVALUATION: NEW PARTNERSHIPS, NEW EVALUATION MODELS
Joy Frechtling, Westat
Donald J. Killeen, Transforming Education Through the Arts Challenge

81 ESTABLISHING CONNECTIONS BETWEEN EVALUATION AND PROGRAM IMPROVEMENT
Denise D. Quigley, University of California at Los Angeles

97 THE PROMISE AND CHALLENGE OF EVALUATING SYSTEMIC REFORM IN AN URBAN DISTRICT
Thomas Corcoran, Consortium for Policy Research in Education
Ellen Foley, Annenberg Institute for School Reform

117 BEYOND STANDARDIZED-TEST SCORES: USING CASE STUDIES TO EVALUATE A REFORM STRATEGY
Ann G. Bessell, Jeanne Shay Schumm & Okhee Lee, University of Miami
Elaine Liftin & Sisty Walsh, South Florida Annenberg Challenge

133 APPENDIX: REPORTS ON THE ANNENBERG CHALLENGE SITES

BEST COPY AVAILABLE

FOREWORD

Warren Simmons and Michael Grady
Annenberg Institute for School Reform at Brown University

RECENT NATIONAL REPORTS and press accounts have documented the full scope of Ambassador Walter Annenberg's unprecedented gift to American public schools. The $500-million Annenberg Challenge provided direct support to locally developed education reform projects in eighteen communities, directly affecting some 2,500 schools. In addition to providing incentives for local schools and systems to improve, the Annenberg gift presented a collateral challenge to education researchers to record and report on the missions and outcomes of the Challenge projects and thus enrich the nation's understanding of the civic, community, and educational conditions needed to promote quality schooling. Each of the Annenberg Challenge projects devoted a share of its grant dollars to support local research and evaluation.

The Challenge Evaluation Studies

The mainstays of the Challenge research effort have been the longitudinal evaluations, designed and conducted by local evaluators working in collaboration with the Challenge projects. Most of these studies were carried out by teams of social scientists from nearby universities or from independent research firms who were involved throughout the life of the project. These systematic assessments of the local reform strategies provided well-specified theories of action, implementation histories, and analyses of student, school, and community-level outcomes. The diffusion of knowledge from these local studies will allow the ultimate impact of the Challenge to be felt well beyond the students, teachers, and communities touched directly by the grants. Information about the final evaluation reports from the sites can be found in the Appendix, along with a list of periodic cross-site reports issued during the life of the Challenge.

Three culminating reports that bring together insights from across the eighteen Challenge sites have appeared in the past year. In June 2002, the Annenberg Foundation and the Annenberg Institute for School Reform jointly published *The Annenberg Challenge: Lessons and Reflections on Public School Reform*, which describes nine lessons learned from the cumulative experience of the eighteen sites. *Lessons* was followed in early 2003 by *The Arts and School Reform: Lessons and Possibilities from the Annenberg Challenge Arts Projects*, offering an analysis and reflections on how the three Challenge Arts projects shaped the character and progress of education reform in their communities.

The collection of essays in this volume is the latest contribution to the body of cross-site knowl-

edge on the Challenge. Our intent here is to further advance the understanding and practice of educators and policy leaders about what it takes to achieve and sustain large-scale reform, all the more urgent in light of new state and federal mandates. In her introduction, Brenda Turnbull articulates the scope of the writing assignment for each of the contributors to this collection, which asked them to address two questions: What did these Challenge projects intend to do and how well did they succeed? How did the partnerships-in-learning between researchers and Challenge program staff at these sites evolve and help shape the work over time?

In Support of Collaborative Learning

This publication appears at a time of great national debate over the fundamental purpose and quality of education research. No Child Left Behind's insistence that improvement strategies be certified through scientific-based research drew strength from recommendations in several national reports, in particular the prestigious National Research Council's *Scientific Research in Education*. That report provides a sweeping commentary on issues concerning research priorities and practices. But what has generated the greatest attention and angst among some scholars in education is its call for expanded federal funding for randomized trials. The fear in some quarters is that federal funding streams for education research will be channeled to support costly, large-scale experiments at the expense of other, equally legitimate forms of inquiry.

We believe that lessons from the Challenge experience bear on the current debate about the state of education research. Fundamental and valuable knowledge was gleaned from the local projects through a thoughtful blending of research methods, always in a nonexperimental context.

In fairness to the Council's report, it states that the essential first step in any scientific inquiry is to identify a research design that suits the empirical question under investigation, and that an experiment is not always the correct choice. That said, we would amplify the caution against a "rush to trials," pressure that might cause government and private funders to reduce support for questions that are vital to our aim of understanding large-scale reform but not easily testable under experimental conditions. Our ultimate aim is not about endorsing one form of inquiry over the other, but rather to challenge the field – both experimental and nonexperimental traditions – to achieve and maintain the highest standards of scholarship.

Achieving and sustaining progress toward the goals of No Child Left Behind will require more and better-quality research on the policy and community contexts of reform. These include questions about effective state and district responses to NCLB mandates, the potential of community partners to advance reform, measures of school quality that supplement standardized-test scores, and the preservation of reform progress in the face of leadership changes. Yet, questions such as these are not effectively studied through randomized trials. Hence, a retreat from these questions by research funders would leave large gaps in our understanding of important aspects of reform in the NCLB context.

The Challenge research experience can also inform a second important proposal in the NRC report, that the U.S. Department of Education's new Institute of Education Sciences adopt design principles that focus its mission and improve operations. One of these recommendations, concerning ways the agency can promote closer working relationships between educators and researchers, is based in the belief that productive relationships between these groups improve the chances that school reform strategies will be designed and implemented using the best available evidence on effective practice.

Most of the projects you will read about in this volume represent strong working models of learning partnerships between researchers and educators that

have yielded greater or lesser benefits for both sides. From the researchers' perspective, a strong, cooperative relationship improves the quality of underlying theories and measures, the proper collection and analysis of data, and the interpretation and consistent application of findings in field settings. By definition, formative assessments are not possible in experiments, given their requirements to operate in more "controlled" research settings.

Likewise, strong collaboration brings benefits from the educators' perspective. One of the main virtues of such a partnership lies in the new access it provides educators to organizations with strong technical capacities to translate research findings from local and national studies into on-the-ground activities and strategies. Scrupulous attention to implementation by researchers affords the project access to data on "leading indicators" of change, information essential to reaching sound decisions about program implementation and deployment of resources.

In closing, we are grateful to the seven local teams of researchers and Challenge project staff who contributed to this volume. Their accounts of empirical discoveries are illuminating; their descriptions of partnerships in social learning, compelling and candid. We are also indebted to Brenda Turnbull and Ullik Rouk of Policy Studies Associates. As editor of this volume, Brenda provided both intellectual and inspirational leadership to this enterprise and wrote the introduction which follows. We hope that the product of their work will help illuminate some of the essential foundations of knowledge needed to sustain Ambassador Annenberg's ultimate goal of investing in public education, which he believed would stimulate "a crusade for the betterment of our country."

References

National Research Council. 2002. *Scientific Research in Education*. R. J. Shavelson and L. Towne (Eds.). Washington, DC: National Academy Press.

INTRODUCTION

Brenda J. Turnbull, Policy Studies Associates

DURING THE 1990s and beyond, the Annenberg Challenge set the standard for ambitious initiatives to reform and support public schooling. Challenge sites in urban centers embarked on reform in response to an invitation issued by Ambassador Walter Annenberg at a White House ceremony in December 1993; nationwide initiatives in arts education and rural school reform did the same.

Across the panorama of sites and projects, much has happened and much has been learned. Some of this learning was shared in *The Annenberg Challenge: Lessons and Reflections on Public School Reform*, a publication that captured significant cross-cutting conclusions about the Challenge and its results (Annenberg Foundation 2002).

This companion volume, *Research Perspectives on School Reform: Lessons from the Annenberg Challenge*, offers a collection of more detailed narratives and analyses from individual sites. It takes advantage of one of the hallmarks of the Annenberg Challenge: in every site a research team worked alongside the reform team, usually in a nontraditional role that involved clarifying plans and advising on next steps, as well as in the conventional evaluation tasks of measuring progress and issuing findings.

As most sites' original funding drew to a close, the Annenberg Institute for School Reform and a cooperating research organization, Policy Studies Associates, invited researchers and site directors to write essays about their experiences and learning in the Challenge. Those who accepted the invitation approached the task from varying vantage points. They included site directors who emphasize what they gained from research and evaluation; researchers who reflect on the puzzles and disappointments that accompanied the sites' accomplishments; and several joint or individual authors who point with pride to their new ways of working. The result is a collection that reflects the rich variety of perspectives from and about research in the Challenge and that begins to assemble discoveries from this pioneering research.

Believing that research and evaluation should inform the work, the Annenberg Challenge encouraged researchers to communicate often and openly with the organizations directing the reforms. The Challenge did not use evaluation as an instrument of central command and control. Instead, sites were free to develop methods, analyses, and reports that would best advance their progress and their learning.

The late Donald Schön played a key role in inspiring many of the researchers to view their designs and working relationships through a "theory of action" perspective. As Schön described it, the researchers would critically observe the sites' intentions

The experiences of Challenge sites present intriguing images of research-reform partnerships suited to ambitious, evolving reform efforts.

(espoused theory), designs (theory in action), and ongoing choices (theory in use) in order to analyze what contributed to successes and failures (Schön & McDonald 1998, pp. 12–13). Organizational learning at the sites would be a major goal of the research.

This volume describes learning that evolved in and from the Annenberg Challenge research. It illustrates a combination of deliberately designed research and evaluation with the adroit seizing of learning opportunities as they arose. Because future reformers and researchers will inevitably face many of the issues and struggles described here, we have sought to record both the conceptual framework and the practical solutions that these authors created. The experiences of these Challenge sites present intriguing images of research-reform partnerships suited to ambitious, evolving reform efforts. They also illustrate ways in which dynamic reform leaders have been able to learn from their experiences, aided by the rigorous inquiry of distinguished researchers.

The Challenge Approach to Reform

Defining its purpose broadly, the Challenge encouraged all the sites to invent reform designs tailored to their settings. The sites represented here illustrate the results of that inventiveness. Moreover, both sites and the schools they supported changed their focus over time as they learned from their reform experience.

- The Center for Arts Education in New York City awarded grants with a philosophy that initially united its cooperating organizations under a banner of mainstream comprehensive arts education. Later evaluation and research led this program into different avenues of funding and inquiry.
- The Houston Annenberg Challenge also supplemented conventional, quantitative measures of results with qualitative inquiry, including hands-on work with individual schools by Planning and Evaluation consultants.
- The Boston Plan for Excellence – Boston Annenberg Challenge worked with the Boston Public Schools, forging an identity as an intermediary organization deeply engaged in school reform, yet independent of the school system in key respects.
- The Transforming Education Through the Arts Challenge (TETAC) blended the approaches to arts education of six collaborating organizations, learned from the evaluation of its initial grants that neither implementation nor results would be easily attained, and changed its strategy in many ways as a result.
- The Los Angeles Annenberg Metropolitan Project launched several initiatives to realize its reform principles. This volume describes the close working relationship between the evaluation team and one initiative, Parents as Learning Partners, which supported schools and networks of schools to involve parents in their children's academic programs.
- In Philadelphia, Children Achieving was the multifaceted reform program of the new superintendent. The project researchers chronicled the results and frustrations of this program and, here, they reanalyze their findings.

- The South Florida Annenberg Challenge relied on case studies to provide in-depth information on reform implementation and about the factors that made a difference in results across different approaches and contexts.

All observers of the Annenberg Challenge agree that its reliance on intermediary organizations in school reform has forged new roles and relationships. The Challenge required that intermediaries – often newly established to lead reform in their sites – differ significantly from school district bureaucracies. They had to be lean, spending a very limited portion of the reform dollars on their own operations. This leanness constrained the organizations' capacity for reflection and learning because they had to do so much with their limited time and dollars. Several chapters in this volume show the mixed results of launching grant programs in haste. The authors also illustrate the struggles in many sites to build working relationships from scratch.

Many Challenge sites tried to be nimble, reassessing and revising their strategies in response to setbacks and new opportunities, and the "theory of action" perspective helped them do this. Site leaders' willingness to learn was a crucial ingredient, and it was often extraordinarily impressive to the researchers, as many of them say in this volume. As Schön had predicted, however, knowing what core principles to hang onto and what tactical choices to abandon was seldom easy. When everything was cast as a matter of principle, as in Philadelphia, the reform was arguably destined to fail because no adjustment was allowed. Other sites illustrated greater willingness to change direction – sometimes frustrating those who saw the changes as compromises.

Civic engagement emerged as a hallmark of many Challenge sites. Researchers in this volume and elsewhere have pointed to hands-on involvement by civic leaders and local philanthropies as a distinctive feature of the Challenge. The drive and commitment of the South Florida Annenberg Challenge board and the engagement of the local Weingart Foundation in shaping the Los Angeles Parents as Learning Partners program are two notable examples.

Civic engagement has meant a coalition of partners at most sites, each bringing a history and agenda to the reforms. Direction could change as partnerships and leadership evolved. Both of the Arts Challenges represented in this volume illustrate this: even though they ostensibly pursued shared aims, the numerous organizations at the table each contributed a somewhat different perspective on engagement of the arts with learning. At worst, this was a problem, bringing confusion where some would have hoped for order. At best, it brought vitality, passion, fresh ideas, and connections to civic assets.

Approaches to Research and Evaluation in the Challenge

Like the reforms themselves, research and evaluation differed across sites, yet shared some common themes. First, all the Challenge projects emphasized student achievement as the goal of reform. In many of the participating communities, philanthropic support for reform had never before insisted on such a level of accountability for student performance. C. T. Kerchner and colleagues (2000) made this point in their summary of the entire Los Angeles Annenberg Metropolitan Project.

In this volume, Denise Quigley echoes the theme of rigorous accountability in her description of the Parents as Learning Partners program within the Los Angeles effort. As Ann Bedell, Jeanne Shay Shumm, Okhee Lee, Elaine Liftin, and Sisty Walsh observe, funders in South Florida wanted the evaluation to "raise the education standards in the community." Reform partners in the two Arts sites were not accustomed

Site leaders' willingness to learn was a crucial ingredient, and it was often extraordinarily impressive to the researchers.

to rigorous outcome evaluation but they learned, over time, to see evaluation as a partner in reform.

Experience in the Challenge has shown that the fate of a reform can hinge on student performance. Tom Corcoran and Ellen Foley describe how Philadelphia staked the credibility of its reform on student gains. Although other factors also contributed to the erosion of support for Children Achieving, both the strengths and the weaknesses of student performance were key to the public dialogue about the fate of the initiative.

The Challenge has supported important work on factors associated with achievement, including research not described in this volume. Research in New York City has demonstrated that small schools are cost-effective when their long-term achievement results are taken into consideration (Stiefel et al. 2000). The Consortium on Chicago School Research has tracked reforms in that city (for which the Annenberg Challenge has provided significant support) and has issued important findings on the connection between authentic, challenging student assignments and achievement (Wenzel et al. 2001). One example of pioneering research on student achievement in this volume is the analysis of Philadelphia data explaining the results of including previously untested students in the accountability system.

A second important feature of research in these sites was that, like the reformers themselves, the research teams had to balance rigorous adherence to their core purposes with constant adjustments to meet changing circumstances. These researchers were nimble in adjusting and even inventing methods. Their creativity went beyond blending qualitative and quantitative methods, although they all changed designs to follow up on promising findings and adapt to new directions in the programs.

Another distinctive feature of these research efforts is the stream of information with which they have rapidly alerted the reformers to emerging results.

Because of their work with the Challenge, evaluators and researchers in arts education have embarked on new advances. In the chapter on TETAC, Joy Frechtling and Donald Killeen describe the balancing of summative and formative evaluation. In some aspects of the design, evaluators took the lead because evaluation integrity had to be paramount, while in others there was more flexibility for joint decision making.

A third distinctive feature of these research efforts is the stream of information with which they have rapidly alerted the reformers to emerging results. This has included providing information on what *wasn't* happening. Early implementation that falls short of reformers' hopes is an old but important story in reform. Frechtling and Killeen call their section on the first wave of TETAC results "There's no 'there' there." The same title could have appeared in other chapters.

Parents as Learning Partners, a relatively small and focused effort of the Los Angeles project, responded to the news of partial implementation by retooling individual participating schools and working vigorously to fill in what was missing. Barbara Neufeld and Ellen Guiney describe how Boston, although a districtwide effort, was also able to make midcourse corrections. Philadelphia provides the opposite case: the central office was largely unwilling or unable to learn from the early news about implementation; researchers could see the "reform overload," but those who suffered from it were unable to fix it.

The researchers also alerted reformers to early successes on which they could build. New York's Center for Arts Education changed its program description from "comprehensive" to "contextual" arts education because the new title better captured the experience at participating schools. Joy Phillips, Pedro Reyes, and Linda Clarke describe what happened in one Houston school when a researcher insisted that the staff ask, "Who are we, and how good are we at who we are?" as a way to focus and strengthen the reform. Early data from the case studies conducted in schools reinforced South Florida's emphasis on leadership.

Finally, some schools participating in the Challenge, like some of the site leaders, experienced research as an ally in their reforms, typically for the first time ever. The Planning and Evaluation consultants' work with individual schools in Houston vividly illustrates several ways in which this happened, ranging from schools' deepening engagement with portfolio evaluation to the "gadfly" relationship that a consultant formed with a principal. In Los Angeles, after the team presented data at individual schools, some principals invited them back for more detailed discussions with school staff and parents.

The Challenge research-reform relationships established levels of trust and cooperation that were truly remarkable.

The Research-Reform Relationship

As should be clear from the foregoing, the relationship between reform and research in most Challenge sites was extraordinarily collaborative. That relationship presents an alternative to conventional images of evaluation as a process that works in isolation to render judgment over the long term. Most of the authors in this volume describe what such collaboration could be like: it had its benefits, but it also brought struggles.

Collaboration was necessary because Challenge-site programs typically presented moving targets. There were several reasons for this: most of the programs themselves reflected multiparty collaborations; they were all too ambitious to be implemented in every particular within a short time; they were buffeted by forces arising from state and local policy contexts; and their leaders made adjustments because they were learning from early results. The Challenge research-reform relationships – most of which were newly forged around the projects – established levels of trust and cooperation that were truly remarkable.

Collaboration was beneficial because it made up-to-date insights available; however, it was never easy. Reformers and researchers live in different cultures and face different professional imperatives. Reformers are rewarded for boundless optimism, big-picture vision, and inspirational images. Researchers' stock-in-trade includes skepticism, careful specification of observed behaviors, and painstaking scrutiny of evidence. The two groups live by different timetables: the Annenberg reformers prized rapid movement; the researchers insisted on time for careful reflection.

Communication about ongoing reforms – with high stakes for the reformers and sometimes-contending reform agendas at the table – is not all sweetness and light. Battles were fought over the meaning of data. As described in *Lessons and Reflections*, every interim Challenge report reflected intense discussions about the wording that would publicly characterize progress and setbacks. Because student testing set a high-stakes hurdle for the reforms, accountability structures within each Challenge site were a contentious issue and sparked intense debate over the merits of the tests.

Implications for Education Research

These lessons are emerging at a time when "scientific" education research is a national imperative and when many are calling for controlled experimentation as the gold standard of scientific inquiry. Clearly, the kinds of research described here are no one's idea of controlled experiments. But the Challenge was by design a dynamically evolving system, and it enlisted research more as a partner than as a judge. The Challenge sites were not installing tightly specified reforms that would hold still for long-term measurement. Rather, they were launching initiatives that

could – and did – evolve. Research informed that evolution in a number of ways documented here. This approach is consistent with two imperatives issued by a National Research Council panel on scientific education research: education researchers should collaborate with education practitioners, and designs must fit the questions posed (National Research Council 2002).

Although the Annenberg Challenge has neither created nor validated simple recipes for reform, the sites' experiences are rich in lessons about both reform and research. The research approaches illustrated in this volume pose important challenges. The sites offer significant examples of success in facing those challenges, but the problems are real and must be faced anew as reform continues in U.S. schools. Future reformers would do well to heed the ups and downs of the reforms chronicled here and would be fortunate to enlist the wisdom of researchers and reform leaders such as those who participated in the Annenberg Challenge.

References

Annenberg Foundation. 2002. *The Annenberg Challenge: Lessons and Reflections on School Reform.* Providence, RI: Brown University, Annenberg Institute for School Reform.

Kerchner, C. T., J. Abbott, D. Ganley, and D. Menefee-Libey. 2000. *The Impact of the Los Angeles Annenberg Metropolitan Project on Public School Reform: An Analysis of LAAMP Board and Staff Activities.* Los Angeles: Claremont Graduate School University, Los Angeles Compact on Evaluation.

National Research Council. 2002. *Scientific Research in Education,* R. J. Shavelson and L. Towne (Eds.). Washington, DC: National Academy Press.

Schön, Donald A., and Joseph P. McDonald. 1998. *Doing What You Mean to Do in School Reform: Theory of Action in the Annenberg Challenge.* Providence, RI: Brown University, Annenberg Institute for School Reform.

Stiefel, Leanna, Robert Berne, Patrice Iatarola, and Norm Fruchter. 2000. "High School Size: Effects on Budgets and Performance in New York City." *Educational Evaluation and Policy Analysis* 22 (Spring), 27–39.

Wenzel, Stacy A., Mark A. Smylie, Penny Bender Sebring et al. 2001. *Development of Chicago Annenberg Schools: 1996–1999.* Chicago: Consortium on Chicago School Research.

CHAPTER 1:
EVOLUTION IN TANDEM: DEVELOPMENT AND RESEARCH IN AN ARTS EDUCATION PROGRAM

Hollis Headrick and Greg McCaslin
The Center for Arts Education

Terry Baker
Education Development Center, Inc.

The New York City Annenberg Challenge for Arts Education

THE CENTER FOR ARTS EDUCATION believes that the arts can make a significant contribution to school change and can improve student performance both in the arts and in the rest of the core curriculum. This belief became the basis for a theory of action that pairs schools with cultural institutions to develop customized approaches to arts education. The Annenberg Challenge supported such partnerships in eighty-two schools in New York City, involving 54,000 of the city's students.

ARTS EDUCATION IN NEW YORK CITY

Two distinct eras define arts education in New York City. The first era is pre-1975. During that time, the city made a financial and programmatic commitment to public arts education that it demonstrated in various ways. Foremost, it offered students across the system opportunities to acquire skills in the fine and applied arts. Students learned visual arts, music, theater, dance, and literary arts from licensed arts specialists working under licensed arts supervisors. The Board of Education had citywide curricula and standards in visual arts and music and some curricula in place for theater, dance, and technical/vocational career studies related to the arts. Its Office for Arts and Cultural Affairs promoted thriving collaborations between public schools and the city's rich cultural community, generally in the form of artist residencies, class visits to museums, and student attendance at performances of various types.

In the mid-seventies, all that changed. Draconian cuts in the city's funding of public schools eroded a previously solid financial and programmatic commitment to arts education. Specialists in visual arts, music, dance, and theater were among the many thousands of teachers laid off. Teachers of the arts who had seniority in the system suddenly found themselves teaching subjects for which they had little or no training or license. Potential arts teachers had no opportunities to teach in their subject areas and either sought jobs elsewhere or turned to other specialties.

The lack of a systemwide means for delivering arts education also meant that schools made little attempt to link the scope and sequence of instruction or accountability mechanisms to instructional standards in the arts. A few pockets of arts instruction survived, but their quality and accessibility were uneven. Arts organizations helped as they could, and in some schools, practicing artists, many of whom had little or no teaching experience, became the primary providers of arts experiences for students.

BEST COPY AVAILABLE

In the early 1990s, the Board of Education resolved that arts education is essential to the basic education of every child and endorsed restoration of arts education for all children.

Renewed Commitment to Arts Education

This state of affairs continued until the early 1990s. By that time, the city's cultural and educational leaders had positioned themselves for what was to become the beginning of a major renaissance in arts education. The president of the school board convened a citywide arts and culture advisory group, which commissioned *Crisis and Opportunity*, a report outlining the dire condition of arts education and its recommendations for changing it. Shortly after that, the chancellor of schools formed a second arts and culture advisory group through the Fund for New York City Public Education (since renamed New Visions). This second group produced the report *A Passion for Excellence*.

In response to these reports, the Board of Education resolved that arts education is essential to the basic education of every child and endorsed restoration of arts education for all children. The board's resolution authorized the chancellor to

- require a policy for arts education;
- provide all students the opportunity to participate in the arts;
- develop procedures to ensure that every school establish arts programs in keeping with standards of the Curriculum Frameworks;
- design an appropriate plan to assess student achievement in the arts and professional development needs;
- maximize opportunities for professional development;
- encourage collaborations among schools, cultural organizations, institutions of higher education, and appropriate community organizations to incorporate the resources of the city's artists, arts and cultural organizations, and institutions of higher education into public schools.

In addition to this action by the Board of Education, foundations and arts and cultural organizations stepped up on behalf of schools and arts education. The DeWitt Wallace–Reader's Digest Fund supported the School Partners Project through the Fund for New York City Public Education. This project became one precursor to the New York City Annenberg Challenge. In addition, an informal Arts Education Funders Group formed. This group, made up of public and private funders interested in arts and education in New York City, provided a forum for discussion on arts education. Its roster of guest speakers included the chancellor of schools and the commissioner of the New York City Department of Cultural Affairs, as well as nationally known practitioners. The chancellor's staff was given responsibility for preparing a list of opportunities for private-sector support of arts education in the public schools. One outcome of this list was support for an Arts Education Resource Center, the antecedent to the Center for Arts Education.

Other collaborations that took place included the New York City Arts in Education Roundtable, an affinity group made up of representatives of cultural organizations that provide services to the city's public schools, the Partnership for After-School Education (PASE), whose members represented community-based organizations, and the Arts and Related Industries Partnership (ARIP), which explored ways to link students with opportunities in the arts and related industries. Arts education was infused with new energy and commitment. When an advisor to Ambassador Annenberg suggested to the commissioner of the Department of Cultural Affairs that New York City establish an Annenberg Challenge for Arts Education, the community was ready.

The Annenberg Proposal/CAE

The New York City Board of Education (BOE), the United Federation of Teachers (UFT), and the New York City Department of Cultural Affairs (DCA) took upon themselves the writing of a proposal to the Annenberg Foundation to restore arts education to the New York City public schools. Underwritten by the Aaron Diamond Foundation, a member of the Arts Education Funders Group, the proposal, "Institutionalizing Arts Education for New York City Public Schools: Educational Improvement and Reform Through the Arts, A Five-Year Plan for Implementation," focused on multiyear grants for schools that brought together teachers, administrators, parents, arts and community organizations, individual artists, and universities.

The intent was to form partnerships to develop sustainable, comprehensive arts education programs. Partnerships would be formed according to a flexible formula joining site-specific needs with available resources at a site. This would allow schools to capitalize on relationships and resources including staff, cultural partners, and community groups. Effective programs were to be strengthened and placed in the context of a school's comprehensive plan for arts in education. Additional programming was to be undertaken as needed. The Center for Arts Education (CAE) would direct the effort.

A press conference in March 1996 announced the New York City Annenberg Challenge for Arts Education. Several large leadership gifts kicked off fund-raising to meet requirements of the Challenge, and an Annenberg Advisory Council was formed, with the commissioner of cultural affairs named temporary chair.

Among the advisory council's first acts was to hire an executive director for the Center for Arts Education and to move into the space in a high school formerly occupied by the Arts Education Resource Center, which had ceased operations. The number of staff hired for the Center quickly grew to four, eventually climbing to eleven three years later when the Center was fully operational.

Starting up any organization is a busy undertaking, and the Center was no exception. An immediate issue it had to address was its own governance. As specified in the plan, the chair of the Advisory Council, the newly designated board chair, and the executive director formed a board of directors and began the process of board development. Members of the board included the chancellor and designee, the commissioner of the Department of Cultural Affairs and designee, the vice-president-at-large of the United Federation of Teachers and designee, representatives of the Arts in Education Roundtable, and leading members of the city's philanthropic, cultural, and business communities.

The Center also needed to create a vision of excellent arts education that would communicate to the city's policy-makers, the Center's potential partners, and the public the core values that the Center would seek to exemplify and support. After much discussion and debate, the board came to consensus on five guiding principles for the Center's work:

- school change through the arts
- arts as part of the core curriculum
- partnership and collaboration
- professional development
- evaluation and assessment

Another task was to contract with the Center for Children and Technology (CCT), a division of Education Development Center in Newton, Massachusetts, to conduct an umbrella evaluation and assessment of the entire partnership program. CCT was to become a major force in shaping the initiative and later reshaping it as midcourse corrections were made.

Meanwhile, the Board of Education showed its commitment to collaboration with the Center. It hired a special assistant for the arts who reported to the deputy chancellor for instruction and professional development. This individual served as the Board of Education's liaison with the Center, attended the Center's activities, and sat in on meetings of its board of directors and program committee.

THE PARTNERSHIP GRANTS PROGRAM

The request for proposals (RFP) for the first of three rounds of partnership grants to schools and cultural organizations was released in December 1997. It was drafted by the executive director, with feedback from the board of directors and advisory group. While the RFP did not prescribe specific practices, strategies, curricula, or outcomes for partnerships, it did challenge them to develop context-specific approaches to the five guiding principles. That is, it asked applicants to capitalize on their assets and deficits in developing their proposals, embodying the guiding principles in ways that honor, reflect, and enhance their unique circumstances. The Board of Education, the United Federation of Teachers, and the Department of Cultural Affairs helped disseminate the RFP and promote the initiative among educators, artists, and cultural organizations. Meanwhile, the Center offered extensive preapplication assistance to schools and cultural organizations as they prepared their partnership plans. In some cases, it helped match appropriate partners; in other cases, it advised proposal writers on technical matters.

Response to the RFP was overwhelming. The Center was deluged with over 400 partnership applications representing virtually all of the arts, from visual arts to performing arts to museum education. Over one-third of New York City's 1,100 public schools and programs applied, forcing the Center to hastily set up a system for managing this huge application load. The Board of Education and the UFT helped manage the intake process, but there was no time to design a programmatic database.

Response to the RFP was overwhelming. The Center was deluged with applications representing virtually all of the arts, from visual arts to performing arts to museum education.

In April 1997, the Center awarded the first round of nineteen planning grants and eighteen implementation grants to thirty-seven schools. Support for these partnerships covered activities from January 1997 to June 2001. Many other partnerships that had applied for grants were disappointed or disenchanted with the results of the first-round process and selection. The Center refined its grants process on the basis of focus groups organized by CCT. Comments and recommendations from school staff, cultural organization representatives, and BOE, DCA, and UFT collaborators helped clarify the five guiding principles and funding criteria and made the grant review process fairer and more transparent. In the second round, the Center awarded twenty-four new partnership grants that supported activities from February 1998 through June 2001. In this round, many representatives from previously funded partnerships served as peer panelists. In a third and final round, twenty-one grants were awarded in 1999 to cover activities for two years, from July 1999 through June 2001. With this round, the scope of the initiative grew to a total of eighty-two schools, 135 cultural and community-based organizations, 54,000 students, and 2,000 teachers participating directly.

Partly in response to the overwhelming interest in the Annenberg arts initiative and the disappointment among schools that were not awarded planning grants in the first round of competition, the Board of Education, with support from the mayor, created Project ARTS (Arts Restoration Throughout the Schools). Project ARTS allocated the first systemwide per capita funds for the arts since the cutbacks of the mid-seventies. The mayor committed $25 million a year for three years. Activities were phased in using a cohort model district by district.

Project ARTS required that community school districts identify an art liaison to oversee the restoration of arts. It also encouraged schools that scored high enough in the peer review process and came reasonably close to being funded by the Center to adapt their "Annenberg Arts" proposals as a framework for restoring arts to their school. Many schools did just this. This was one of the earliest contributions of the Annenberg Arts to the system as a whole.

Implementation Issues

Several issues impeded implementation. For example, Center staff had to determine what the five guiding principles looked like in action. This involved close collaboration with the partnerships that were implementing the principles in schools and much discussion with other stakeholders, both local and national. Establishing and maintaining relationships among diverse players became a major undertaking. Sometimes this meant building bridges where previously there had been none; at other times, it meant mending fences to keep collaborations moving forward intact. Different stakeholders had different goals, priorities, values, levels of commitment, and expectations for the Center.

For example, some stakeholders viewed the goal as transforming education, reshaping schools, and redefining learning. Others wanted to restore arts instruction in a more traditional way by hiring more arts specialists. Some educators expected cultural organizations to offer professional development that would equip classroom teachers with the skills and confidence to integrate the arts into their instructional strategies and thereby ensure arts restoration beyond the grant period. Still other stakeholders wanted to become real partners with their schools, active in shaping school policy and practice, or in developing new financial resources for schools. Some stakeholders envisioned a school system that defined itself more by the range and diversity of student experiences and accomplishments than by performance on standardized tests.

One reason for such disparities was that key collaborators in the field had little time to develop a shared understanding of the initiative's guiding principles and what they meant for schools. The level of their understanding varied enormously, particularly among schools and cultural organizations that had little experience in integrating the arts into the school's program. The lack of time for partnerships to develop a shared vision also interfered with team building, as different visions pulled team members in different directions. This was compounded by district and citywide leaders' focus on literacy and high-stakes testing, which eclipsed and in some cases displaced the restoration of the arts, frustrating some school-based teams who had to compete for teachers' professional development time. And, of course, all of this occurred in an environment of constant turnover as principals, partnership contacts, cultural organization representatives, and district superintendents left one position for another, leaving Center staff to develop new relationships with their replacements.

Establishing and maintaining relationships among diverse players became a major undertaking. Sometimes this meant building bridges; at other times, it meant mending fences.

Some of these issues resolved themselves over time as partners worked together to implement their plans. But others required significant adjustment. For example, the district required that teachers participate in professional development on literacy and test preparation. Finding time on top of that for professional development and partnership planning in the arts was difficult. Some schools found that they had proposed to do more than they could, or that they didn't have the space for it all. Similarly, the varying levels of skill and commitment among teaching artists and cultural organizations signaled the need to intensify professional development. Delays in payment of per session fees for planning and professional development undermined teachers' morale and dampened enthusiasm for change through the arts.

Staff from the Center dedicated significant time to visiting schools, meeting with new partnership members, and building leadership. They also supported

The alignment and coordination between partnership activities and evaluation activities required constant attention, negotiation, and technical assistance.

and encouraged periodic gatherings where representatives from schools and arts organizations could form loose networks. Peer-to-peer problem solving addressed logistics, communication, evaluation, and assessment. Center staff also convened regular meetings of representatives of the Board of Education, the United Federation of Teachers, and the Department of Cultural Affairs to review the initiative's progress.

Staff daily struggled with balancing the need to stay in the office with the need to be out seeing and working with the partnerships on the task of making the ambitious five-year plan operational. Each new round of grants brought in an enormous load of applications. While this response demonstrated the overwhelming interest on the part of schools and partnerships in restoring the arts to schools, managing the process was a daunting undertaking for staff.

Nor did the deluge stop once partnerships were funded. Project coordination was a constant challenge. Not unexpectedly, funded partnerships needed many kinds of help, including how to write proposals, how to design professional development and implementation plans, and how to assess their partnership's impact. Partnerships also had questions about the long-term effects and durability of the initiative, especially when up against skepticism about its efficacy in the field at large.

Evaluation, too, raised issues for partnerships that had to be resolved. Partnerships did not always understand the purposes of evaluation, and the kinds of student assessment requested by the schools and the CAE program were new to many of the artists. Cultural organizations were accustomed to evaluation of their programs, but not of the impact of their programs on students. The program raised evaluation and assessment standards to new levels. Its two-level design called for participants to provide their own local evaluations to assess student learning in the school and to participate in a larger program evaluation conducted by an outside agency. To do this, however, it had to raise the capacities of adult participants to conduct and become better consumers of evaluation. The alignment and coordination between partnership activities and evaluation activities such as surveys, site visits, and interviews required constant attention, negotiation, and technical assistance.

Evaluation Strategy

In the Center's early planning, the notions of comprehensive arts education and partnerships combined to form a theory of action. According to this theory of action, a number of elements work together to enhance students' mental, emotional, and social growth and contributions to society, eventually enriching economic, intellectual, and community welfare. These elements consist of the arts themselves, their associated skills, aesthetic contexts, integration into the education of the young, and the unique ways that effective teachers and teaching artists present them to students in the context of learning and understanding. Partnerships are the catalysts that bring these elements together and spur local school reform.

Five key research questions guided CCT's evaluation of the Center, the partnerships, and the arts education provided through the Annenberg Challenge:

- In what ways is the nature of arts learning qualitatively different when outside cultural resources partner with schools to design/deliver curriculum?
- How does the integration of the arts support school-change efforts?

- In what ways is student learning in non-arts subjects improved through the introduction of the arts? How is student learning in arts and non-arts subjects improved through partnerships?
- Do the arts provoke parent and community involvement in a school, and in what ways is this linked to school change?
- What is the impact and legacy of this sustained partnership of local cultural organizations, in terms of their capacity, understanding, and experience working with schools?

A Two-Pronged Effort

The partnership program had a two-part evaluation strategy. First, the Center required local partnerships to conduct their own studies, using external evaluators or local partnership participants, to determine the impact of their partnership activities on the school and students. CCT reviewed these evaluations with an eye toward

- locally generated assessment plans
- student-achievement data
- student cognitive development data
- student attitude data
- student arts performance data
- technical assistance sessions on assessment for schools, artists, and arts organizations
- the variety of assessment techniques employed by local partnerships

These annual evaluations focused on local school programs, not citywide issues or "outcomes." As such, they provided a great deal of information to schools and the Center about the structural and instructional content of the local school efforts, but they did not provide the kinds of information originally intended about student performance and program impact.

The second part of the evaluation strategy involved CCT's evaluation of the entire partnership program. From the very beginning, a close collaboration between the Center and CCT was a design characteristic of the evaluation approach. Following Robert Stake's "responsive evaluation" approach (Stake 1975) and the "design research" approach of Allan Collins (1990, 1993) and other staff at CCT, the Center adopted an intense form of formative evaluation that would provide timely and ongoing assessment of the program as it was implemented. This would allow the program or implementation strategy to be adjusted, or "formed," in process rather than waiting for final, or "summative," judgments about effectiveness. The evaluators' belief, and that of the theorists they followed, was that it made more sense to try to correct the course and work toward success than to simply make final judgments when it was too late to make corrections. But to do such work required a close collaboration built on trust.

Prior to beginning its evaluation, CCT reviewed program design documents, such as the five-year plan, the initial RFP, and the first round of planning and implementation grant applications. It also observed selection panels and readers in their deliberations. At the request of the Center, CCT hosted meetings and focus groups that informed program planning and allowed the assessment team to collect statements of attitude and opinions about the initiatives from leading arts educators, artists, and arts organization administrators. It administered a needs-assessment survey to grant recipients and observed planning-grant recipients as they devised activities. As evaluation plans developed, CCT reviewed the local evaluations and assimilated their results into a final program evaluation.

Center staff and the CCT evaluation team met each quarter to review and, wherever possible, align program implementation and research. Alignment and realignment, in particular, were constants in the

The Center adopted an intense form of formative evaluation that would provide timely and ongoing assessment of the program as it was implemented.

program, which had always been defined as evolving. The local programs asked for clarification of what the Center wanted in its evaluations. CCT worked with the Center staff to provide cross-site sharing sessions in which evaluators compared their products, CCT staff explained their reading of the local projects' reports, and local project staff and evaluators received technical support in methods and reporting. In addition, the research director at CCT began attending Annenberg Challenge cross-site gatherings with CAE representatives. From these meetings, they received useful information on how the theory of action model was supposed to work and began developing instruments to share with others through the Annenberg Institute for School Reform.

From "Comprehensive" to "Contextual" Arts Education

Initially, Center-funded partnerships were assessed to determine

- the extent to which they affected the whole school and every child;
- the extent to which they worked toward common school reform goals;
- the impact of their parent involvement and co-learning activities;
- the ways they provided for staff planning and professional development;
- their assessment process and the result thereof;
- their plans for expansion during and sustainability after the funding period.

The adjustment from a program described as "comprehensive" to one described as "contextual," a direct response to CCT researchers' findings, required providing additional technical assistance and adjusting the evaluation design.

As local partnerships began developing their programs, it became evident that their instructional designs fit no one pattern. The "comprehensive arts education" model was replaced by a more accurate depiction of the actual instructional practice, captured by the phrase "contextual arts education." CCT adjusted some of its evaluation strategies to better ground them in actual practice. For example, CCT evaluators shifted from attempting to locate curriculum programs that were consistent across schools to accounting for individualistic instructional strategies. They shifted from studying "curriculum art" with clear and consistent scope and sequence to documenting what they came to call contextual arts education that varied from site to site, depending on the context defined by local resources, themes, topics, access to cultural organizations, and core curriculum alignments at each site.

The adjustment from a program described as "comprehensive" to one described as "contextual," a direct response to CCT researchers' findings, required providing additional technical assistance and adjusting the evaluation design. In making this shift, CCT had to reconsider its basic theory and methods of analysis. It had to determine "what counts" in contextual instruction, how to count it, and how it could be explained or transferred to other schools. If context is, by definition, idiosyncratic, what lessons can be drawn from contextual instruction that would have relevance to the general population or to other schools? These questions forced adjustments in CCT practices, and their answers are still being investigated. It had been relatively easy to justify a comprehensive or curriculum focus by thinking about its applicability to other sites and about the dissemination of effective practice to other schools or cultural organizations. Contextual programs, in the arts or other curriculum disciplines such as math or technology, require different ends and different justifications, and thus the analysis has to change as well. This ensured more accurate documentation and more trustworthy theory development than work from hypotheses that were fixed and generalizable.

The initial choice of a comprehensive and sequential curriculum model for the project echoed the wishes of many educators in the United States who have called for a standardized curriculum across

school districts. But the reality of education in the American context is, in Howard Gardner's terms, "highly dispersed." Gardner (1996, p. 104) reports: "'Context' has not been my favorite concept, but I have gained a new respect for its importance." Although he is referring to in-school curriculum arts, the contextual nature of the work is yet more complex when partnerships around arts in education programs are developed by schools and cultural organizations jointly. It has become increasingly important, as partnership programs have expanded with renewed funding, to account for context in the assessments of student learning and the evaluation of instructional programs.

The last quarter century has witnessed a sea change in basic conceptions of how learning occurs. Neurology, anthropology, and psychology provided new evidence on how the human brain works and how social and cultural contexts provide necessary linkages for thought and learning. Contextual understanding emerges from knowing and learning through shared activities and experiences and helps define knowing and learning as "synonymous with changes in the ways that an individual participates in social practices" (Cobb & Bowers 1999, p. 6). Such thinking is taking hold in the psychology and education research communities, stimulating new research and provoking new debates about learning and instruction.[1]

The ways in which contextual variables are incorporated into instructional design and evaluated by researchers have become the defining elements in measures of success. Measures of achievement, impact, or implementation that do not attend to complex sets of variables are incomplete. Just as it is important to design arts education around those characteristics of the arts and arts experiences that are necessary for their definition, so is it important to evaluate arts education programs according to contextual variables. If such programs "must create a new context," then research and evaluation efforts must attempt to document and account for the ways in which the new contexts are shaped by the programs. Such research

1. Key studies and reports that identify points of connection or linkages in student learning and development learning and identify understanding as the ultimate outcome include Lave 1988, Gardner 1983, Collins et al. 1989, Pea & Kurland 1986, Caine & Caine 1994, and Resnick 1987.

The ways in which contextual variables are incorporated into instructional design and evaluated by researchers have become the defining elements in measures of success.

should, as Winner and Hetland (2000, p. 6) say, "explore the ways in which the arts may change the entire atmosphere of a school. This way we can begin to understand how the arts affect the 'culture of learning' in a school. We can then develop rich, qualitative measures to evaluate whether the arts lead to deepened understanding of – and engagement in – non-arts areas." CCT's evaluation work aimed to create rich documentation of context variables in an arts education partnership and the ways that students, schools, and communities change in response to new combinations of variables (Baker 2001, p. 6).

The focus of evaluation was now on the extent to which

- a partnership provides adequate instructional time, content in the arts, participatory learning activities, and interdisciplinary studies;
- this instruction is attractive and engaging for students;
- students are learning by doing;
- students acquire a broad range of abilities and knowledge within the specific disciplines they study;
- instruction is developmentally appropriate;
- the program changes student attitudes toward art.

A Focus on Leadership

The following year (1998–1999), evaluators put the spotlight on school partnerships, with special emphasis on their leadership aspects. Given that the partnerships' approach to leadership was bottom-up with top-down support, evaluators considered local school leadership the most important level to examine. But they also examined shared leadership among the

initiative's main partner organizations, the Center, the Board of Education, the United Federation of Teachers, and the Department of Cultural Affairs, as well as each organization's independent leadership role.

Partnerships were assessed to determine

- whether or not the partnership formed its own local committee;
- who the members of the local partnership committees were;
- whether or not the committee held regular meetings;
- which partner organization played an "anchor" role;
- the content of year-end reports.

The Center itself was examined to determine the extent to which it

- built bridges between the educational system and external providers of arts education;
- held constituencies accountable to each other;
- channeled the efforts of the artistic community/industries;
- resisted budget constraints and political shifts;
- initiated and supervised the proposal development process;
- provided pre- and post application technical assistance to applicants;
- oversaw review of proposals according to standards;
- provided assistance to schools/districts that had not worked with external partners;
- provided citywide professional development, leadership sessions, national model sessions, demonstrations, and presentations by educational and artistic leaders;
- facilitated selection of exemplary models and best practices.

Working with Partnerships

Relationships between partnerships and the Center continued to develop and, in some cases, deepen. Good relationships were characterized by regular communication, active participation in Center-sponsored events, and response to the Center's requests for information and documentation. Partnerships in good standing with the Center took part in funder visits, presentations at Annenberg cross-site meetings, and panels at other professional gatherings. In some cases, representatives from partner organizations served as peer group facilitators, hosted visitors, and spoke at conferences.

Staff at the Center worked as closely with partnerships as seemed comfortable. In several cases Center staff participated in partnership planning meetings and became an active resource to the partnership by highlighting strategies that were effective in other contexts, by acting as a sounding board or by identifying potential resources, financial and otherwise. Center staff continued to visit partnerships to observe workshops, planning meetings, parent events, and other activities and to facilitate next steps when necessary. One of their main jobs was to maintain high expectations without being prescriptive as the partnerships put legs under their visions.

Classroom practices were varied, as the contextual approach suggests, so a single example or even set of examples of good practice does not convey the total impact of the partnership program. An example of a program from a participating high school is illustrative.

The program at this high school consisted of six different year-long "arts studios" co-taught by a teacher and a teaching artist. Students were placed by grade level in a studio of their choice. Each week throughout the year, they attended a two-hour art studio class designed to develop their arts skills in a given domain (acting, dance, visual arts, videography, design, poetry). This was one of the more intensive and sustained of the CAE partnership programs in terms of student contact hours with the arts and also in terms of professional development for teachers and artists.

Skills, Sequence, and Arts Integration

At first glance, assessing the development of arts skills would seem to be the obvious approach for assessing project impact. Many of the students at this

school, however, particularly in the program's first two years, started with very little exposure to and experience with the arts. Assessing their development of arts skills with objective high school – level indicators would not necessarily be appropriate.

There is also the issue of arts integration. The project's initial goals of integrating the arts with non-arts areas were altered to "linking" the arts to the non-arts. In theater classes, for example, in the first year of the project, ninth- and tenth-graders wrote and performed plays around the idea of imperialism, the theme of the humanities curriculum for that year. In the second year of the project, the program was changed so that tenth-graders read and performed plays from the World War II period, the focus in their humanities courses, with an emphasis not on the play's content, but on the reading and performance of the play.

The content links were thus made more oblique but were intended to be mutually reinforcing. And, in fact, in the theater course where scenes from *The Diary of Anne Frank* were being rehearsed, a researcher observed the humanities teacher discussing with students the Nazi occupation of the Netherlands when students began to consider the stage sets for scenes from the play. Disagreement about the size of the stage attic space led to discussions about how and why Jews were hidden in the homes of the Dutch. In the exchange, the teaching artist who had been leading the class faded into the background as the humanities teacher led the discussion. After a while, when the students seemed satisfied with the conversation, the teaching artist took over again to move the rehearsal along.

At the same time as direct links to the curriculum faded, there was also no sequential development of arts skills across grade levels, or even within a grade. Teaching artists developed arts skills rubrics, but they were not seeking to move each student along a continuum of development. The overall program, which allowed students to change studio arts class each year or to stick with the same one, did not differentiate between novices and more experienced students and was not structured for the sequential development of essential skills. Instead, the program was to a degree "product-oriented" (with a balancing emphasis on "process"), with periodic panels of outside practicing artists coming in to provide critical feedback to student performances or exhibitions. Through this and other project components, the program successfully connected a relatively isolated group of high schoolers – economically and socially – with the arts community. It built local community support for the school, including funding alliances.

Changes in how students perceived and related to their community, and especially the cultural community and to the arts as a cultural force in society, were worth examining.

The accomplishment of which the project administration spoke most highly was the extent to which the arts programs came to "matter" to the students in the school. "Students now see the arts as something that is their right," said one administrator. Teachers, too, were beginning to demand participation in the program, she reported. Art and "culture" became a central feature in the whole-school curriculum. Teachers were asking that the arts become part of their regular weekly planning meetings.

Documenting the Impact of Partnerships

What did all this mean for assessing impact? An objective assessment of arts skills alone might only prove successful for more naturally talented students, given the lack of a sequential approach to learning. An assessment of learning in non-arts areas, such as world history, would be difficult (and perhaps meaningless) to link to the arts. But the changes in how students perceived and related to their community, and especially the cultural community and to the arts as a cultural force in society, were worth examining. How the school changed through its new alliances with local community representatives was likewise

Maintaining respect for practitioners' time and vision while pressing them for time to address critical research needs was a constant conflict and worry.

interesting to examine. In this case, there was also substantial change on the part of the cultural organization, which came to understand the world of schools and classrooms in entirely new ways. Working with our theoretical framework of how context-rich partnership programs can change school culture, we documented the complex connections and relationships that contributed to these changes.

Maintaining respect for practitioners' time and vision while pressing them for time to address critical research needs was a constant conflict and worry. Simply by its existence, the project raised issues having to do with evaluation and assessment, documentation, managing change, intrapartnership communication, shared decision making, integrating arts curriculum with an eye to scope and sequence, and planning to sustain the partnership after the initial grant period.

The Center convened its first cross-site gathering in May 1998. Teams from sixty-one schools and 100 cultural organizations attended. The purpose of this meeting was to permit individuals participating in the initiative to meet each other, exchange ideas and experiences, and help Center staff prepare for future activities. Participants met by peer group within each borough – teachers met with other teachers, artists with artists, and parents with parents – to discuss successes and challenges. Center staff took notes for follow-up action. Feedback from this conference led to further meetings among specialty groups. For example, a one-day conference for teaching artists and arts organization representatives, called Developing a Common Language, was held in June 1998 to highlight the need for artists and arts organizations to negotiate educational issues including standards, child development, school reform, school logistics, and so on. Educators were not invited to this conference in deference to their knowledge of these issues and their many end-of-the-school-year responsibilities. Another meeting drew local partnership evaluators to discuss their evaluation plans and challenges. Out of this meeting came Compelling Evidence, a one-day conference on evaluation tools and methods held in October 1998 with the research team from CCT.

Three more such cross-site gatherings were held, making give-and-take a major characteristic of the initiative's first phase and helping develop a sense of community across partnerships. A network began to take shape and extended to CCT. The Center held two gatherings for partnership representatives to review initiativewide formative evaluation findings with CCT's research director.

While these activities were full of useful findings for the initiative, there was clearly a need for additional data on the impact of the arts-infused curriculum on teaching and learning. To this end, CCT and the Center designed and piloted a practitioner action research project Student Learning in and through the Arts. This project, funded by the National Endowment for the Arts and General Electric Corporation, is still in progress. It funds teams of teaching artists and classroom teachers to refine student assessment in their classrooms, while CCT evaluators document their activities and coach them on the design of research strategies. Teaching artists and teachers meet together to plan, develop, test, refine, and assess fully integrated arts curricula in this project. CCT will continue to document these processes and results including

- diverse models of arts-integrated curriculum with embedded student assessments,
- documentation of the curriculum development process, illuminating how choices were made, how specific arts domain strengths were drawn upon, and how student learning was monitored,

- documentation of student learning in the relevant arts and non-arts domains;
- analysis of the curricula that illuminates connections to learning standards and other systemwide curricular frameworks.

This work expands on the work done by the CCT team and local evaluators. It adds to the quantity and quality of investigation in the crucial topic of student performance assessments in arts education, helping to establish validity and depth.

In another series of sessions, teachers and teaching artists came together to reflect on their practices in looking at student work. At another series, they learned about effective ways to use video for documentation and developed their own examples of video documentation.

Interagency Efforts

Interagency efforts to foster arts education among such organizations as the New York City Board of Education, the Annenberg Challenge, the United Federation of Teachers, and the Empire State Partnership Program of the New York State Council on the Arts/New York State Education Department challenged Center staff to listen, learn, and lead. Throughout the implementation and evaluation of its own efforts, the Center collaborated with these other organizations to form and strengthen a collective focus on arts education. The Center mounted a major campaign to build support for arts education and to coordinate its efforts with intersecting and overlapping systems. Staff spoke at a staggering number of conferences and other events. At the request of the Open Society Institute, which was in the process of establishing the After-School Corporation, staff discussed arts in the after-school setting. The Center participated in the citywide Arts Education Week, sponsored by the Board of Education and community cultural leaders, and it went to the Empire State Partnership Summer Seminar, which focused on professional development and peer exchange.

In January 1999, the Center hosted its own large-scale public event. Promising Practices: The Arts and School Improvement was a conference highlighting practices developed by the partnerships. Peer sharing sessions were organized according to the Center's five guiding principles (school change through the arts, arts as part of the core curriculum, partnership and collaboration, professional development, and evaluation and assessment). Partnerships proposed sessions and critiqued each other's session ideas in planning meetings facilitated by Center staff. Attendance exceeded 500 at this meeting.

A major development in the Center's collaboration with other organizations was the design of a career development program in the arts. Designed to create opportunities for high school students and educators to participate in school-to-career activities in the arts and related industries, the program grew out of a study commissioned by the Center. The pilot Career Development Program got off the ground with the recruitment of a director and the development of an evaluation plan by CCT's research team. The Center recruited high schools, teachers, students, and job sites to participate in this pilot course of study and internships. It also convened a Career Development Advisory Group made up of representatives from labor, the school system, the for-profit and not-for-profit cultural communities, and higher education. With findings from the pilot program and practitioner feedback, the Center refined and expanded the Career Development Program.

More than 325 students and fifty work sites have participated in the program since it began in 1999. These students and educators have gained valuable experiences in the arts and arts-related industries and have explored their individual interests and learning about career opportunities. Most students in the program have attended college, and 80 percent of student interns have continued to pursue careers in the arts and arts-related industries, from fashion design to journalism. Several students have either been hired or have extended their internships as a result of their participation in the program in organizations such as Teachers & Writers Collaborative, ABC, Ballet Hispanico, Kenneth Cole Productions, International Center of Photography, and Nola Recording Studios. Relationships have been estab-

lished with two union work sites, providing student interns with technical training and access to the union trades, which can be difficult to access.

Other activities were designed to generate interest in the initiative as well. The United Federation of Teachers offered to highlight the work taking place by underwriting a publication on school improvement through the arts at Center-funded partnerships. Center staff identified a writer and a designer and created a preliminary outline for the publication, "Promising Practices: The Arts and School Improvement."

The development of a Center Web site extended this sense of community even further.

THE IMPACT OF THE CENTER'S ARTS PROGRAMS

Since its inception, the Center for Arts Education (CAE) has contended that adding the arts as content to the school program constitutes a significant school reform effort. According to CAE, arts instruction can improve student performance, both in the arts themselves and in the rest of the core curriculum; make a significant contribution to school change at several levels; foster parent and community involvement in the schools; and develop capacity in community organizations. Our findings supported this theory.

As described above in the sections Implementation Issues and Evaluation Strategy, measuring the arts education program's impact called for new approaches to evaluation. These new approaches would need to provide feedback about student performance but they also needed to go beyond looking at traditional student-achievement data to consider the broader impact on students and schools and to provide timely feedback to the partnership. CAE used a two-level evaluation design. First, local partnerships conducted their own studies to evaluate the impact of the program on student learning and on the school, using external evaluators or local partnership participants. Second, local partnerships participated in the larger program evaluation conducted by CCT.

To varying degrees, the evaluation activities at both levels included analysis of quantitative data such as student achievement and staffing levels, along with data about perceptions, attitudes, new activities, and changes in practice from surveys; observations during visits to classrooms, workshops, and meetings; interviews; and compilations of written reports from practitioners. The evaluation reports from the local sites were examined by the CCT team to compare results, methods, and data with CCT's focus-school documentation data to compare results and to discover areas that were not covered in either effort. As a result, new evaluation responsibilities were assumed by the CCT team, especially in the area of student impact. In 2000 and again in 2001, CCT, working with the research office of the New York City Board of Education, identified standardized-test data in reading and math that was available for the third- and fifth-grade student populations at the time and conducted two separate analyses of those data to determine impact of the program on student learning in core curriculum areas. Documentation and description of practices were an important part of the evaluation, in addition to quantitative performance evaluation.

Although effective assessment models for arts programs are still in development, our evaluation yielded some clear positive results, which are summarized in this section.

School Change

Arts education programs had a noticeable impact on Center schools at the level of curricula, staffing, instruction, and teacher professional development.

Curricula and School Staffing

The CAE partnership had an important impact on student access to arts instruction in Center schools. The number of school arts staff doubled between 1996 and 2001. More students received sequential arts instruction in all arts areas (50 percent more than in 1995–1996, the year prior to the start of the partnership grant program, used as the baseline year from which change was measured). With the addition of the CAE Career Development Program, twice as many students in CAE partnership schools received career preparation as in 1995–1996.

Some schools reported the development of a "distributed leadership" model where teachers throughout the school took on responsibility for the programs. Some schools hired additional arts staff to

work with the teaching artists of the CAE partnerships program. Project coordinators judged those programs to be most successful in which certified arts teaching staff were integrated into the project. Prior experience with other arts programs before the CAE partnerships was highly correlated to the project coordinators' perception of effectiveness in assessing student progress, gaining higher student achievement, and delivering more skilled instruction.

Integration with the core curriculum, in areas such as social studies, history, English language arts, mathematics, and science, was the most frequently observed approach to arts instruction. The nature of arts integration varied from project to project and from classroom to classroom. It also varied with the capacity of teachers and teaching artists.

Some examples of arts skills instruction were seen throughout the program sites that followed this approach; the local evaluation reports contain some brief descriptions of teaching artists' practices. However, with the exception of the few certified arts teachers connected with the program, the evaluation team did not see classroom teachers concentrating on arts skills instruction, nor do the local evaluation reports feature such practices. More commonly, the teaching artists taught the arts skills required for the use of a particular art form in integrated instruction lessons rather than teaching arts skills developmentally or sequentially.

A long-term presence of teaching artists in the classroom seemed to be more effective than a short-term presence. The more time teaching artists spent teaching with their partner, the more they thought that working with the teacher benefited classroom practice (effect size 0.69) and that students were buying into the project (effect size 0.89), and the more cultural organization administrators thought that the role of the arts was enhanced in the school (effect size 1.08). This is a clear finding in favor of more intensive, prolonged arts residencies, proving them to be more effective in injecting the arts into the school.

Teachers and teaching artists did not often use the phrase "aesthetic education" to describe their instructional or curriculum approach in the CAE program – only one site specifically identified itself with the aesthetic education philosophy, though several others indicated that they emphasize aesthetics.

Instructional Practice

Principals indicated in their interviews that changing teachers' instructional practice was their primary goal and expectation for the partnership program. The responses of teachers and teaching artists indicate that this expectation was met successfully.

Partnerships between teachers and artists changed the nature of instructional delivery. Teaching artists and arts organizations learned about the New York State Learning Standards and standardized-testing requirements and developed new ways to support the implementation of standards in classrooms in areas linked to state reading, math, and Regents' tests. Teachers co-taught with teaching artists. Some teachers actively co-designed and taught the integrated lessons, thus developing new abilities to collaborate and co-teach. Others played more passive roles in the classroom, as observers or sometimes as disciplinarians.

Teachers were exposed to a wide variety of community resources, from materials brought in by teaching artists, to work with agencies new to them, to new roles developed for parents. Teachers came to use new methods of evaluating student progress and learning. One common claim of teachers and administrators is that the arts programs allowed them to see students in a new light. Teachers incorporated arts activities into their instruction when the teaching artist was not present. Likewise, teachers used new classroom management techniques acquired from teaching artists.

Teaching artists reported that they experienced significant changes in their own practices – listening more carefully to the needs of teachers, looking for

One common claim of teachers and administrators is that the arts programs allowed them to see students in a new light.

The process of co-planning and co-teaching was itself a source of professional development for both teachers and teaching artists.

curricular connections, thinking about student learning and assessment, and learning more about developmentally appropriate instruction. The teaching artists came into the classroom as professional artists – experts in their fields – bringing passion and knowledge about their arts domains and introducing students and teachers to new role models and ways of being in the world.

Professional Development

Professional development offerings to teachers varied widely from project to project. On average, projects reported nine professional development sessions per year. New teachers were inducted into the culture of the school and practice of arts integration through professional development activities.

CAE offered an ongoing series of gatherings called Looking at Student Work. A total of forty-one partnership projects participated, sending fifty-four teaching artists, forty-three teachers, and evaluators to attend one of three groups of eight sessions where they discussed student art work and the kinds of learning they found in the student work. Student Learning in and through the Arts invited ten teams of artists and teachers to work with researchers to document their arts-integrated lessons. Although the project was designed simply to capture and describe in some detail the nature and effects of the arts integration lessons, it unfolded as a professional development project for teachers and artists.

Partnerships reported success in cosponsored professional development opportunities for teachers and teaching artists sharing their respective expertise in areas such as classroom management and arts pedagogy. Successful teacher professional development emphasized direct interaction with the art form, helping to overcome barriers of teacher fear and inexperience. Teachers indicated that they had increased their knowledge about art forms through professional development focused on the art forms being taught in their schools. These sessions often were modeled on the types of classes the teaching artists would teach for the students. However, some teachers reported not having time to take advantage of professional development activities because they were required to participate in other district or BOE-mandated professional development in math and literacy.

District or BOE mandates for professional development in other areas often inhibited partnerships' abilities to focus on the arts in learning opportunities for teachers. Competing mandates from the districts forced schools to make choices about their limited professional development time. The amount of time that the programs could devote to the arts or arts integration or assessment was not large. The process of co-planning and co-teaching was itself a source of professional development for both teachers and teaching artists and has led to increased understanding on the part of cultural organizations of the demands placed on schools and increased exposure on the part of teachers to the means and modes of instruction in the arts.

Student Learning

The Education Development Center/Center for Children and Technology (EDC/CCT) team collected reports of student learning from the local site annual evaluation reports that were sometimes substantiated and sometimes not. That principals, teachers, and teaching artists were convinced of the power of the learning experiences that the arts provided is not in doubt. However, the systemic capacity for practitioners to frame questions and gather evidence and to analyze that evidence so that substantial statements can be made about student learning was extremely low.

Projects raised the question: "What are students learning?" But few partnerships developed the expertise to implement assessments that captured and usefully analyzed student learning. Many teachers and principals felt that standardized-test data

were not the best place to look for any substantiation of a powerful and engaging curriculum and student learning. Teachers and principals rely on many more indicators – such as student engagement, attendance, and behavior; the connections students draw between lessons; and the quality of student work produced in the classroom.

Many of the judgments that teachers and principals make, and the ways they reach them, remain undocumented. There was an increasing tendency from 1998 to 2001 for the local evaluation reports to cite student learning of arts skills (69 to 86 percent), learning non-arts content (31 to 66 percent), appreciation of the arts (23 to 37 percent), expanded creativity and imagination (23 to 42 percent) and achievement of standards (20 to 34 percent). During the same period, evaluation reports increasingly (15 to 24 percent) noted improvement in reading test scores, a finding that the EDC/CCT analysis of Board of Education reading test scores supports. Project evaluations provided rich information about the kinds of experiences provided for students, but gave a less vivid image of what students gained from these experiences.

EDC/CCT conducted an analysis of a stratified sample of New York City (NYC) standardized English Language Arts (ELA) test scores. Twenty-four Center-funded schools were identified as target schools for analysis. The following summaries compare target schools with other public schools in the same socio-economic status (SES) category. The comparison is based on percentage of students meeting the fifth-grade NYC ELA requirement (reaching levels 3 and 4 in the exam).[2]

- The mean percent of students meeting the requirement in our target schools for 1999–2001 was 40.1. The mean percent of students meeting the requirement within similar NYC schools was 36.3. This was a total difference of 3.8 percent; that is, each of the target schools, on average, was located 3.8 percent above the general NYC school performance for 1999–2001. This difference is not strong enough to conclude that Center-funded schools distinguished themselves from the general NYC school performance.
- When breaking down the number by years, the mean difference in 1999 was 6.7 percent; in 2000, 3.3 percent, and in 2001, 1.5 percent. These findings are also not strong enough to draw conclusions and do not support our theory of accumulating impact, according to which we would have expected an upward trend from 1999 to 2001.
- Fourteen (58 percent) of the target schools were located above the NYC mean, and ten (42 percent) were located below it. While this information is positive, it still is not large enough to establish cause or to support our expectations.
- The twenty-four schools include seventeen schools from low SES groups (groups 7–12), and seven from high SES groups (1–6). Interestingly, six out of the seven (86 percent) high-SES schools are located above the NYC mean, while only eight out of the seventeen (47 percent) low-SES schools are located above the NYC mean. This finding may indicate that the CAE funding raises performance mostly for high-SES schools, and less so for low-SES schools.

Altogether, the partnership schools did not differ greatly from the expected mean of NYC schools. Though the trend is in a positive direction, when looking at the entire sample, the favorable trend is too weak for us to conclude that the CAE funding has affected student performance on standardized tests. However, when looking at high-SES schools alone, the improvement is evident.

In 2000–2001, we looked at long-term-funded schools and at fifth-grade scores that were not appropriate for the earlier study. The data we analyzed in 1999–2000 contained only the 1997–1998 school-year data (the most recent data available when we did the study in 1999–2000). These data were drawn from the school year after the arts partnerships had been funded for one or two years. In 2000–2001, we analyzed 1999–2001 data (after the arts partnerships has been funded for four or five years). There were many changes in the configuration of the pro-

2. Our findings do not derive from the study of data at the same schools. Therefore, our 2000–2001 findings neither support nor contradict the conclusions we drew from the 1999–2000 findings.

gram at the school level – participant population, grade level, curriculum, cultural organization affiliation – that prohibited seeing cumulative impact. Therefore, the results were more likely to show the cumulative impact only of those years of treatment and to differ from the data in our first analysis.

Parent Involvement, Community Partnerships, and Systemic Capacity

A major aim of the CAE partnership program was to expand the involvement of parents and community in arts education. In addition to stimulating ties to parents and the community, the partnerships had a large impact on the capacity of local cultural organizations. The CAE partnership also made significant progress in capacity at the system level. These systemic changes were distributed through the various components of the program and show up in data on schools, cultural organizations, and the program itself.

Parent and Community Involvement

With the Department of Cultural Affairs, the Center offered grants of up to $5,000 to 204 schools for a Parents as Arts Partners program to educate parents about the value of the arts in their children's education and encourage parent advocates. This project serves 22,000 parents annually. The CAE Career Development Program provided orientation, training, and fifteen-week internships for students from high schools at almost forty work sites.

CAE and EDC/CCT collaborated on the development and implementation of a research effort funded by the National Endowment for the Arts on Student Learning in and through the Arts, supporting teams of teachers and teaching artists as they document, assess, and describe the student learning and achievement that occurs when an arts-integrated curriculum is taught. In partnership with the United Federation of Teachers, CAE produced *Promising Practices: The Arts and School Improvement* (Marrapodi 2000). CAE distributed 1,100 copies to public schools, district arts liaisons, local politicians, major contributors, and over 200 cultural organizations. The large demand called for a reprint of the publication.

CAE established and operates a gallery at 180 Maiden Lane in Lower Manhattan to present student art work from participating schools, with three rotating exhibitions managed by Center staff. CAE's advocacy and communications office, with sponsorship from PaineWebber Incorporated, produced a "4Rs" public awareness campaign to focus public attention on the arts as an essential component of a child's education. The campaign included mass-transit advertising, a full-time hotline service that received more than a thousand calls, information packets, and a special subsite on the CAE Web site. CAE staff and members of the evaluation team extended the program's influence by participating in the Learning Partnerships meetings of the Arts Education Partnership (a national arts education advocacy organization in Washington, D.C.) and in documentation efforts at the national level.

Partnerships with Local Cultural Organizations

CAE partnerships made important contributions to the capacity of local cultural organizations. Cultural organizations (COs) gained access to new funding sources. They hired new staff for arts partnerships programs and created new types of positions to support partnerships, such as project managers and coordinators. CO administrators said their organizations had changed the way they develop curriculum and programs.

COs began to work in arts disciplines that were new for them, adding, for example, dance, visual arts, and music to their historical repertoire. COs began to address education reform issues such as learning standards and student assessments, many for the first time in their institutional histories. COs changed their curricular focus even in projects outside the

A major aim of the CAE partnership program was to expand the involvement of parents and community in arts education.

scope of the Arts Partnership program. COs reported integrating their arts curriculum with core curriculum areas for the first time.

COs reported that they were forming new partnerships with schools outside the CAE partnerships program and using their curriculum and teaching methods developed in the CAE partnerships to work with those schools. CO administrators said they had changed the way they evaluated work of teaching artists and that they had changed the way they provided planning time to practitioners such as teachers and school administrators. Partnerships established structures for communication and planning to ensure that basic program obligations were met. The arts partnerships usually, but not always, overcame the service-delivery (or vendor) model that many teachers and teaching artists were accustomed to.

Systemic Capacity

The CAE partnership program also developed significant capacity at the system level, including planning and managing activities across agencies and organizing professional development and technical assistance across the program.

CAE helped coordinate New York City arts education efforts and planning by creating bimonthly management update meetings of leaders from the Board of Education, the United Federation of Teachers, the Department of Cultural Affairs, and the Center for Arts Education. In cooperation with DCA, CAE participated in several policy and advocacy efforts with the mayor's office, influencing the creation with the Board of Education of Project ARTS for all public schools in New York City.

CAE conducted preapplication and technical assistance workshops for potential project sites and followed funding with Starting Smart sessions on issues and expectations regarding evaluation and assessment and budget and finance. CAE staff participated in and assisted the Empire State Partnerships Project in its summer seminar professional development series. CAE conducted twenty-four sessions, each consisting of eight three-hour workshops, on looking at student work in which teaching artists and teachers from forty-one partnership projects participated. Staff development workshops were provided by CAE for 147 members of local project teams and some guests on resource development and proposal writing. CAE partnerships participated in an intervisitation program for 111 participants, including teaching artists, teachers, school and cultural organization administrators, evaluation staff, and a team from the Minneapolis Arts for Academic Achievement program, who visited five local school projects.

Cultural organizations began to address education reform issues such as learning standards and student assessments, many for the first time in their institutional histories.

CAE conducted four annual cross-site gatherings for 1,475 Center-funded project staff from both schools and cultural organizations to discuss partnership issues such as evaluation, curriculum, leadership, and sustainability. CAE designed and conducted a citywide gathering focused on developing a common language for school and cultural organization personnel. CAE conducted a citywide convocation of evaluators and project staff to explore what constitutes and how to collect "compelling evidence." In collaboration with Project ARTS and the Council of Supervisors and Administrators, CAE sponsored a school leadership conference for principals and district personnel on sustaining change and conducted a second Institute on Using Cultural Institutions as Instructional Resources in August 2001. CAE and the partnership schools and organizations strengthened their

links with citywide support efforts such as the Arts Education Roundtable and shared their work through Roundtable workshop sessions.

In collaboration with the EDC/CCT evaluation team, CAE supported a series of four three-hour meetings in an evaluators' exchange series for independent partnership project evaluators. CAE and the EDC/CCT evaluation team conducted an Implications for Action session for all project personnel to review the evaluation report and to explore ways that evaluation can be a tool for program development.

THE FUTURE

CAE and CCT are now embarking on a new cycle of program development and implementation. Building on the research and evaluation that guided its initial five years, CAE will concentrate on disseminating and demonstrating its promising strategies, practices, and programs.

CCT will also continue to evaluate the impact of arts infusion on school and classroom cultures, the working nature of partnerships, and the leadership and management issues involved. In addition, it is planning small, tightly focused research studies of the impact of a controlled set of crucial variables. The purpose of these studies is to draw out knowledge that is embedded in instructional practice. By keeping the studies small, researchers expect to more nearly approximate real learning conditions and settings, making the results more useful to teachers and teaching artists.

For example, CCT will continue to research systemic, reliable, and valid means for measuring the effects of arts learning on students' cognitive, social, and personal development. Researchers will develop a set of assessment instruments designed to measure arts learning, academic achievement through the arts, and cognitive and social development through the arts.

References

Baker, Terry L. 2001. "Contextual Arts Education Assessment: Counting in Context." Paper presented at the annual meeting of the American Educational Research Association, Seattle.

Caine, Renata, and Geoffrey Caine. 1994. *Making Connections: Teaching and the Human Brain*. New York: Addison-Wesley.

Cobb, Paul, and Janet Bowers. 1999. "Cognitive and Situated Learning Perspectives in Theory and Practice," *Educational Researcher* 28: 2 (March).

Collins, Allan. 1990. *Toward a Design Science of Education*. Technical Report no. 1, Center for Technology in Education. New York: Bank Street College of Education.

Collins, Allan. 1993. *Design Issues for Learning Environments*. Technical Report no. 27, Center for Technology in Education. New York: Bank Street College of Education.

Collins, Allan, John Seely Brown, and Susan E. Newman. 1989. "Cognitive Apprenticeship: Teaching the Craft of Reading, Writing, and Mathematics." In *Cognition and Instruction: Issues and Agendas*, edited by L. B. Resnick. Hillsdale, NJ: Lawrence Erlbaum.

Gardner, Howard. 1996. "Rejoinder to Steers." In *Evaluating and Assessing the Visual Arts in Education*, edited by Doug Boughton, Elliot W. Eisner, and Johan Ligtvoet. New York: Teachers College Press.

Gardner, Howard. 1983. *Frames of Mind: The Theory of Multiple Intelligences*. New York: Basic Books.

Lave, Jean. 1988. *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. New York: Cambridge University Press.

Marrapodi, Maryann. 2000. *Promising Practices: The Arts and School Improvement*. New York: Oxford Lithograph. Available at <www.cae-nyc.org>.

Pea, Roy, and D. M. Kurland. 1986. "On the Cognitive Effects of Learning Computer Programming." In *Mirrors of Mind*, edited by R. D. Pea and K. Sheingold. Norwood, NJ: Ablex.

Resnick, Lauren. 1987. *Education and Learning to Think*. Washington, DC: National Academy Press.

Stake, Robert. 1975. *Evaluating the Arts in Education: A Responsive Approach*. Columbus, OH: Charles E. Merrill.

Winner, Ellen, and Lois Hetland. 2000. "The Arts in Education: Evaluating the Evidence for a Causal Link," *Journal of Aesthetic Education* 34: 3–4 (Fall/Winter).

CHAPTER 2:
BUILDING CONSTRUCTIVE PARTNERSHIPS IN URBAN SCHOOL REFORM

Joy Phillips and Pedro Reyes
University of Texas at Austin

Linda Clarke, Houston Annenberg Challenge

The Houston Annenberg Challenge

DURING THE BEGINNING years of the Houston Annenberg Challenge, constructive partnerships (Schön & McDonald 1998) helped tremendously in the design and refinement of this reform initiative for public schools. Schools teamed up with evaluators in innovative ways to reflect on their reforms and improve their effectiveness. This chapter describes the context of the reform; the philosophy of the evaluation team; examples of constructive partnerships; unanticipated outcomes; and the lessons learned in this process.

"AN ACADEMICALLY RICH AND PURPOSEFUL EDUCATION"

The Houston Annenberg Challenge was launched in the greater Houston area in 1996 through the collaboration of several individuals from local foundations, educational institutions, and corporations. These key individuals expressed concern about the quality of public education in the greater Houston area and sought to develop an organization capable of initiating and nurturing systemic change. They committed to conducting a "multidistrict, citywide campaign that focused the community's energies, political will, and financial resources on a strategic investment in networks of public schools that with their community partners would thoughtfully work toward whole-school change" (Child-Centered Schools Initiative 1996).

The broad-based community group, led by representatives from the Brown Foundation and Houston Endowment Incorporated, created a vision for the public school reform initiative and in March 1996 formed a nonprofit organization named the Child-Centered Schools Initiative (CCSI) of the Greater Houston Area. The mission of this newly formed organization was "to promote an academically rich and purposeful education for more of our children and to demonstrate how such an education could become possible for all children" (CCSI 1996).

Because the accomplishment of this mission would depend on a major infusion of public and private dollars, CCSI prepared a proposal for funding from the national Annenberg Challenge. The proposal addressed three key issues: teacher learning, reducing schools' isolation, and size (that is, creating personalized learning environments for children).

The vision for the Houston Annenberg reform effort became a reality with a one-to-two matching grant from the Annenberg Foundation, which contributed $20 million with the stipulation that Houston raise $40 million in public and private matching funds. This funding led to the creation of the Houston Annenberg Challenge.

The CCSI planners designed the Houston Annenberg Challenge work using the national Annenberg model of school reform and lessons learned from other Challenge sites.[1] They focused on building each school's capacity to promote an academically rich and purposeful education for more of Houston's children. The planners focused their vision of local reform at the school level on the three key issues of isolation, size, and, in particular, teacher learning:

> This reform will confront directly the role of the teacher as key to the education of the children. . . . In short, the distinctive component of the Houston Child-Centered Schools Initiative that sets it apart from and makes it a valuable model for reform in other cities will be its broad commitment to provision for teacher learning. (CCSI 1996)

Our research on the reform confirms that teacher learning did, indeed, become the centerpiece of the Houston reform as the planners had intended (Reyes & Phillips 2001).

While the three key issues identified in the original proposal remained at the core of the Houston work for the entire five years of the Houston Annenberg Challenge, initiative planners also envisioned including parents and community members in the reform program. The proposal identified potential community partners such as universities, community organizations, cultural institutions, and corporations. To foster a culture of inclusion, planners invited participation by representatives of dozens of community organizations including the arts, health care, children's services, neighborhood organizations, parents' groups, social service agencies, and private businesses. Furthermore, the planners proposed establishing a regional faculty to provide technical assistance to reforming schools in an effort to reduce isolation among schools and to help sustain reforms generated during the Houston Annenberg Challenge. The regional faculty concept was modeled after the National School Reform Faculty.[2]

A Commitment to Collaboration

In many ways, initiative planners recognized the critical role of collaboration in the work that lay ahead. The original proposal frequently mentioned the collaborative nature of the reform plan; collaboration was identified as a crucial process, especially with regard to addressing two of the three key issues: breaking down isolation within and among schools and between schools and communities, and dealing with issues of size by personalizing the learning environment. Planners most likely anticipated that creating collaborative relationships across multiple sectors in Houston – a city known for diverse interests and strong political differences – would not be easy. Our research confirmed the difficulty of that task, particularly in the later stages of planning and in the early stages of implementation.

Despite the emphasis on collaboration, we suspect that even the planners did not anticipate the richness of the collaborative relationships that would emerge among teachers, between teachers and administrators, and between practitioners and community-based researchers. While many of these collaborations began as attempts to implement strategies to reduce isolation and personalize the learning environment, our data reveal that collaboration among partners developed into a key mechanism in teacher learning. Furthermore, administrators' and teachers' efforts to reduce size and personalize the learning environment often grew out of collaborative professional development activities, such as site visits to model programs. We believe that these "collateral effects" began with the evaluation philosophy adopted by educators and external evaluators.

1. The original proposal from Houston to the Annenberg Foundation reported that preliminary planning work began in 1995 through a newly formed 501(c)(3) community organization called the Child-Centered Schools Initiative. This organization began doing business as the Houston Annenberg Challenge after being named one of the eighteen national Annenberg Challenge projects. At the time of this writing, the organization was selecting a new public name to reflect its post-Challenge work.

2. The National School Reform Faculty was created in 1995 by the Annenberg Institute for School Reform at Brown University. In July 2000, it relocated to the Harmony School Education Center in Bloomington, Indiana. For more information, see <*www.harmonyschool.org/nsrf/default.html*>.

A Commitment to Evaluation

One of the most important aspects of these collaborations was the philosophy of evaluation. The overall theory of action was simple, yet elegant. School administrators and teachers would construct the particular theory of action that would guide the education reform at their school. At the same time, they would work with the external evaluators to design strategies and activities. Thus, the first step in the reform implementation process was to build a common perspective on the reform among all stakeholders.

External evaluators made the evaluation process less intimidating for school staff by using formative feedback as the preferred strategy to help implement the theory of action. Formative feedback meant communication between evaluators and schools around strategic and performance plans, evaluation activities, feedback to the school in the form of performance reports, and program improvement activities. Formative feedback required a commitment to methodological pluralism, using both quantitative and qualitative methods to obtain information. In addition, evaluators focused on person-referenced evaluation and participatory action research, which meant that all stakeholders participated in the evaluation design and data collection. Finally, evaluators emphasized decision-oriented knowledge and context-specific results – that is, findings that the schools could use.

Evaluation of the Houston Annenberg Challenge

The Houston Annenberg Challenge Board of Trustees commissioned a three-year (1999–2002) independent evaluation to be conducted by researchers from the University of Texas at Austin, the University of Houston, and Rice University. The purposes of the research and evaluation study included determining

- how the funded schools put the reform initiative in place;
- what the schools did as a result of the initiative;
- what apparent impact the initiative had upon schools and teacher learning;
- what apparent impact the initiative had upon students' academic performance.

In our evaluation, we drew upon multiple sources of qualitative and quantitative data to assess progress of reform efforts, using two major strategies: a macroanalysis of all funded schools and a microanalysis of a subset of schools.

At the macro level, we compared student test data from Annenberg-funded schools with the academic performance of each school's own students in prior years and with the academic performance of comparable peers at other schools. Also, we designed a set of surveys for administrators, principals, teachers, students, parents, and community members across all funded schools.

At the micro level, we searched for evidence of campus changes in teaching and learning. We selected twelve schools – six elementary, three middle, and three high schools – for intensive case study. We considered the reform's impact in three broad areas: student outcomes, school development and teacher learning, and the building of support for systemic change.

Student Outcomes

- Academic achievement in reading and mathematics as assessed by using the Texas Assessment of Academic Skills test, reported in the Texas Learning Index, to compare campus-level performance longitudinally within and across schools (Texas Education Agency 2003)
- Alternative assessment of student work – including portfolios, anthologies, artwork, oral defenses, and performances – to evaluate depth and breadth of student knowledge
- Survey data collected from students, teachers, parents, and school partners to describe stakeholders' perceptions

School Development and Teacher Learning

- Campus-level changes in teaching and learning
- Relationships within and among funded schools
- School organizational structure

Building Support for Systemic Change

- Sustainable impact on funded schools
- Observable impact on greater Houston area

BEST COPY AVAILABLE

We believe the process of building cohesion among stakeholders in the evaluation program helped tremendously in negotiating the political tensions typically associated with evaluations. Many times participants in school reform are threatened by the evaluation process and, as a result, build resistance to change. In this case, however, teachers, administrators, parents, and members of the community worked together to design evaluations, collect data, and make sense of data with the external evaluator. Therefore, the evaluation philosophy helped to implement and successfully modify the theory of action.

The Houston Annenberg Challenge made a commitment to assess and modify its initiative through evaluation at several levels: participation in a national cross-site analysis of Annenberg projects, comprehensive evaluation of the Houston initiative, and network- and school-level evaluation. When the first cohort of schools was selected for funding and formed into a network called the Beacon schools, the Houston Annenberg Challenge required these schools to include evaluations and documentation of their work in their plans.

The significant investment in network- and school-level evaluation proved valuable to many of the reforming schools, and it is the primary focus of this chapter. These evaluations were intended to provide accountability for funds and to assess student performance and progress toward reform goals. Additionally, the evaluations were intended to reduce school isolation by increasing collaboration among schools, within schools, and between schools and the community. For instance, principals, teachers, and parents were invited to make site visits to other schools to observe classes, interview staff and children, review children's work, and examine the schools' documentation of reform efforts. The planners encouraged schools and networks to use both standardized and nonstandardized data collection instruments.

The evaluation philosophy helped to implement and successfully modify the theory of action.

Funded schools were expected to design comprehensive evaluation plans to measure progress toward the three key issues – improving teacher learning, restructuring all aspects of school size, and reducing isolation – as well as providing evidence of increased student learning. To help with this formidable task, reforming schools were allowed to recruit external evaluators who could provide technical assistance and assist the schools with evaluation design and implementation. Using these guidelines, Houston reformers created a mechanism unique among the national Annenberg projects: the Planning and Evaluation consultant.

CREATING CONSTRUCTIVE PARTNERSHIPS

Applying their philosophy of evaluation to school reform in the greater Houston area, reform organizers created several mechanisms for constructive partnerships between school insiders and outside researchers. Schön (1983, 1987) and Argyris and Schön (1992) have described a constructive partnership as a type of formative research and evaluation conducted between an initiative's insiders and outsiders who have been invited to explore the inside of the initiative.

The Planning and Evaluation Consultants

During the early days of the initiative implementation, organizers decided to require each of the eleven Beacon schools to use part of its Annenberg funding to contract with university- or community-based researchers. These researchers, hired as Planning and Evaluation consultants, quickly became known within the reform as P & Es. Because each consultant perceived the role differently, the consultants' services to funded schools varied considerably. In this section, we highlight examples from the work of three consultants. These individuals, according to many sources, contributed positively to their schools' reform work during the five-year initiative.

Some participants believe the concept of the P & E consultant was borrowed from the Chicago Annenberg Challenge. The Chicago project used the

concept of external partners as a framework for a potentially broad set of relationships between schools and community organizations. According to the Chicago plan (Newmann & Sconzert 2000), possible partner functions included

- helping schools develop curriculum, instruction, and assessment techniques;
- providing and structuring professional development opportunities;
- providing and facilitating leadership development opportunities for local school councils, parents, and community members;
- brokering other outside resources;
- providing coalition or networking support;
- organizing community involvement;
- advocating on behalf of the schools.

All participants interviewed agreed that implementing the P & E consultant process was an unstructured, although not chaotic, course of action. Initially, the eleven Beacon schools held half-day organizational and content meetings every month. The principal, the school's Annenberg grant coordinator, and the P & E usually attended together.

Different visions surfaced immediately. Some of the initiative planners advocated for a more traditional, structured approach to implementation, documentation, and evaluation of the reform process. Practitioners from the funded Beacon schools resisted this approach as excessively top-down and formed a coalition to lobby for a school-based approach. Beacon school practitioners argued that their schools had been chosen because they were already engaged in innovative reform strategies, and that they were therefore capable of figuring out the best way to capture the new Annenberg-funded work. The process, they thought, was formative; they would invent new strategies as they went along.

As the P & E consultants began working with the funded schools, they put the idea of creating a constructive partnership between administrators and teachers into practice. Schön and McDonald (1998, p. 13), building on the work of Argyris and Schön (1992), explain how a constructive partnership differs from traditional evaluation.

> If the insiders are an initiative's architects and actors, and the outsiders its evaluators, then the constructive partnership implies an evaluation methodology quite unlike the common variety, in which evaluators work in a relatively "hands-off" way and seek to objectify causal connections between program interventions and their outcomes. . . . It does not seek to hold "treatments" stable but to subject them to continuous reflection. And its boundaries are the boundaries of the action situation it studies, leaving open the relationship between what it may discover in the situation and what might be discovered in others. Illuminating this relationship requires additional inquiry and reflection, or what we call reflective transfer.

If the insiders are an initiative's architects and actors, and the outsiders its evaluators, then the constructive partnership implies an evaluation methodology quite unlike the common variety.

Shortly after the initiative formally began in the spring of 1998, the first Reforming Schools Summer Institute introduced the theory of action concept to administrators, teachers, and P & Es. Helping practitioners within reforming schools to create a theory of action and guiding them to reflect continuously about their work became a primary mission of the P & Es. Schön and McDonald (1998) define theory of action as an analytic tool used in evaluation "to help practitioners (including designers and implementers of reform) reflect upon and make explicit the knowledge that shapes what they do; in other

words, it is a tool that helps them inquire into and learn from their own practice" (pp. 10–11). Reform implementers always construct a theory of action; they can never borrow one from another setting because it must be based upon each individual school context.

A theory of action has three phases: what reformers say they would like to do, what plans they design to help them achieve the goals they've set, and what actions they actually take. Schön and McDonald describe these three phases or facets of the concept as espoused theory, design theory, and theory in use. Ideally, reform implementers demonstrate coherence among the three dimensions; but, in reality, reform work ranges along a continuum from little coherence to a great deal.

The theory of action approach became the framework around which the funded schools created action plans for the reform. It was used as a way to document the work of the initial years of funding, generally in evaluation documents that the P & E consultants wrote for their schools at the end of the first and second funding years.

The P & E consultants created unique constructive partnerships with their school partners, the nature of which depended upon their skills and strengths. One of the three consultants was particularly interested in teacher knowledge and school context and was well grounded in the use of qualitative methodology. The other two, trained in educational psychology and statistics, drew on their more traditional measurement and evaluation backgrounds. However, all three shared a willingness to work without a superimposed evaluation structure, preferring, instead, to design with their school partners a localized process based on the school's context. All seemed comfortable embarking on a journey of exploration with their schools without any preconceived notions about what lay ahead.

Reform implementers always construct a theory of action; they can never borrow one from another setting because it must be based upon each individual school context.

All three P & Es believed strongly that part of their role was to connect their school partners to relevant academic literature. The consultants believed that by helping practitioners reflect upon their own tacit knowledge while connecting them with current literature, they could facilitate conversations that would help the practitioners move their schools forward. Many P & Es also became deeply involved in a second reform mechanism: accountability and peer review.

The Accountability and Peer Review Process

Another important mechanism of the Houston reform was the accountability and peer review process. Although initiative planners believed strongly that accountability strategies were essential, they disagreed on how such strategies should be designed and implemented. Some planners favored creating a traditional, structured evaluation system for the funded schools to follow. Others advocated for school-level autonomy, allowing practitioners at each funded school to create independent evaluation strategies. Principals at the eleven initially funded schools – the Beacon schools – objected strongly to the idea of being given a template. These principals, who characterized themselves as having very strong personalities, argued, "We were chosen to be these top [Beacon] schools because we have some great ideas of what we want to do and how we want to do it.... We are more than just a collection of standards."

Out of these lively discussions came the idea of each school telling its own story (Craig 1997, 1999). For several schools, telling their own story meant developing school portfolios. For administrators at one elementary school, an Educational Toolkit became both the reform strategy and the evaluation mechanism. The founders of an experimental high school used their ideas about authentic assessment to design strategies for tracking the reform progress. Faculty

from each school, working with their P & Es, ultimately created evaluation models that took into account each distinctive school context.

In 1998, the fledgling Houston Annenberg Challenge created the School Accountability Report (SAR), a reporting document to help the newly funded schools implement accountability measures. One staff member recalled developing the first SAR.

> That very first year we decided we needed a more definitive rubric for the schools to use, because if we were asking them to change their performance, we had to have some way to measure the change. This rubric began to form the basis for what we were looking for in the schools. The report [SAR] has three elements that we ask the schools to demonstrate: partnerships, leadership, and sustainability. These three elements also incorporate the three imperatives of improving teacher learning, reducing size, and reducing isolation.

Thus, the reporting format asked schools to detail their progress on each of the three imperatives through partnerships, leadership, and efforts to make change sustainable. In addition to describing specific actions and outcomes, each school rated its activities on a continuum of the reform process, from "beginning" through "emerging" and "systematic" to "sustaining."

To introduce the SAR format and the concept of theory of action, the intermediary organization sponsored a training event, the Reforming Schools Summer Institute (RSSI). The intermediary staff used annual RSSIs and periodic Action Labs to provide intensive training about reform to funded and nonfunded schools in the greater Houston area. For example, one Action Lab focused on the use of evidence in reform documentation. According to the objective for the training, "educators will acquire a deeper knowledge of monitoring educational outcomes by collecting, analyzing, and applying a variety of data to make strategic decisions to improve teaching and learning." Staff used documentation ideas from the Coalition of Essential Schools such as common and uncommon measures (Cushman 1996). They also incorporated a "cycle of inquiry" approach to continuous improvement, adapting the approach taken by the Annenberg-supported Bay Area School Reform Collaborative.

Faculty from each school, working with their Planning and Evaluation consultants, ultimately created evaluation models that took into account each distinctive school context.

Another strong component of the Houston Annenberg Challenge accountability system was the peer review process, which was linked closely to the SAR documentation. Peer review teams consisting of two teachers, one administrator, and one community person (e.g., a P & E consultant, parent, or business representative) brought people together across districts. The teams conducted daylong site visits using the school's SAR documentation as their guide. They used Critical Friends Group protocols,[3] including warm and cool feedback, to structure their visits. Team members could do walkabouts, interview faculty, and examine student work.

A principal participating in the initiative's Principals Academy reflected on a site visit:

> The opportunity to tour the school in a large group and also the opportunity to tour in a smaller group focusing on a particular question worked really well for me. As a result of this experience, I will use this format in my Learning Community to gain vital information on how we can improve student engagement.

Intermediary staff recalled that early site visits were informal – designed, for the most part, for looking at model programs. Gradually, as the site visits

3. Critical Friends Group (CFG) is a form of teacher-led study group that originated in 1995 at the Annenberg Institute for School Reform as a program of the National School Reform Faculty.

became connected to the accountability system, the visits became more structured. However, participants stressed that the visits were always about learning and sharing information rather than monitoring compliance. As one staff member observed, "It's as if the peer review process was just a huge, giant P & E."

The Case Study Researchers

A third mechanism of collaboration to emerge from the Houston Annenberg Challenge was the close working relationship that developed between case study researchers[4] and the practitioners at the case study schools. Like the P & E relationships, these relationships varied according to the individual. We spotlight two relationships between case study researchers and practitioners in this section: one in a middle school and one in an elementary school.

The middle school case study researcher joined the project in the second year of the study. She described her data-gathering role as "all about establishing relationships." Since she considered the principal's support crucial, she began by spending quite a bit of time with the principal, "getting to know each other." This researcher believed that establishing trust was essential to gaining access to the classroom teachers. As a former middle school science teacher herself, the researcher bonded immediately with the principal. Always on the lookout for good teachers, the principal even offered the researcher a teaching job "on the spot."

The graduate student assisting this case researcher expressed concern at the onset about proceeding without a structured research plan that included interview and observation protocols. Believing in the formative nature of the process, the lead researcher recalled, "She was worried and asked me, 'What are we going to do?' I was comfortable in replying, 'I don't know. Let's just go talk with the principal and see what develops.'"

Shortly after beginning her work at the school, this researcher was invited to give a keynote presentation at a district-sponsored conference on implementing project-based learning. Many teachers from the middle school attended the conference. The researcher reported that, after the presentation, the teachers' perception of her changed "from one of them [an outsider from the university] to one of us [an insider, a teacher]."

As the teachers became more comfortable with her, they were willing to open up and talk about reform implementation. She saw her role not only as a data collector, but also as a resource person. As she observed teachers in their classrooms and met with them individually, she shared relevant information about curriculum sources such as Web sites. Some teachers began to see her as a mentor. This researcher described the case study relationship as ultimately "very personal." For the formative process to be effective, she believed, practitioners must trust the researcher's motives.

A member of another research team conducted his study at an elementary school ethnographically. For the second year of the study, he literally moved into the community where the school was located. A former early childhood teacher with a strong literacy background, he worked closely with the principal and teachers. He helped guide implementation of literacy-focused professional development, tutored in an after-school reading program, and substituted in classrooms.

This researcher was especially interested in studying teacher identity and teacher knowledge in the context of school-based reform. He believed that establishing a trusting relationship was key to understanding bigger issues of how localized school-based reform strategies interact with systemic reform strategies, such as state and district accountability measures. In this collaborative relationship, the researcher became a quasi member of the faculty.

4. University-based researchers were asked to conduct case studies in twelve funded schools. These schools – elementary, middle, and high schools – included Beacon and Lamplighter schools. Generally, teams of one university faculty person and one graduate student carried out the case studies.

COLLATERAL EFFECTS AND OUTCOMES

The mechanisms discussed in the previous section provided the infrastructure that helped school reform take root in the schools. The work facilitated by P & Es, case researchers, and others involved in the accountability process helped create some unexpected outcomes. In this section, we describe these outcomes, which we call "collateral effects" of the reform.

Portfolio Group: Writing Outside the Lines

One P & E was asked by five Beacon schools to serve as their consultant. Initially shocked at the idea, this P & E decided she would consider it only if the schools' representatives agreed to work together as a group. The practitioners – a collection of principals and teachers – agreed. The group coalesced around their shared conviction that they had the ability, at the school level, to create effective methods for documenting and evaluating their work. This collaborative group of Beacon practitioners came up with the idea of school portfolios.

Initially, these school portfolios resembled coffee-table scrapbooks. One or two people at each school typically constructed the first portfolios out of artifacts and photographs from school events, achievements, and, occasionally, student work. The practitioners saw these early portfolios as devices to justify their Annenberg funding and to enhance other grant applications.

Over the five years of their work, however, teachers from the five Beacon schools[5] watched their process and product change. Gradually, their portfolios began to capture more of the complexity of school change. The portfolios increasingly included teacher reflections, student growth, and individual voices. Teachers said, "As our buildings are being reconstructed, so are we."

More and more faculty became involved in the process of building the annual portfolio, and in the process, the documents moved from artifacts to voices. The portfolio group summarized their collective insights about the value and validity of the process in which they had been involved at a symposium presentation at the annual meeting of the American Educational Research Association in New Orleans in April 2000.

Initially, the school portfolios resembled coffee-table scrapbooks. Gradually, however, they began to capture more of the complexity of school change, increasingly including teacher reflections, student growth, and individual voices.

Some teachers were involved with the portfolio group from the outset. Others joined as the process unfolded. Still others moved in and out of the process. One member of the last group described his intermittent participation as an expression of frustration. A middle school language arts teacher, he had initially resisted the idea of journal writing as too structured. Working with one student, he had an insight that "you don't have to write on the lines." His student had a habit of submitting assignments on scraps of paper, fast food bags, or napkins. Despite the unorthodox materials, this teacher recognized growth in his student's writing as the boy moved from "entries full of anger with very little elaboration" to "descriptive, well-organized essays." As this teacher began to understand the power of individual style, he also began to relax in his own writing. Gradually, his own journal entries "outside the lines" revealed

5 An additional school eventually joined the first five schools in the portfolio group. This school, an Annenberg-funded middle school, is a member of the Lamplighter school network. Teachers at this middle school recognized the value in the portfolio process and asked to join the group. Although funded at a different level and therefore in a different phase of reform, this school was welcomed by the other five into the portfolio group and has presented publicly with them about their work on numerous occasions.

The teacher team concluded that portfolio development created value for individual teachers, groups of teachers, and the entire school.

increasing understanding of his students and the effects of his pedagogy on his students' work.

> Based on his journal, I allowed him to sit by himself and write independently if he chose. I think he feels safe that way. He is always first to class, and I have tried to speak to him and encourage him daily. He really seems to focus and take constructive criticism well now. (And I have to do less and less encouraging and more and more discussion of his increasing strengths as a writer.)

This teacher described the school portfolio and his journal as dance partners in the writing process. As a member of the school portfolio team, he agreed to keep a reflective journal. This journal became part of his personal teacher portfolio. Portfolio team members kept journals in a variety of styles and forms. Most faculty used their portfolio experience to guide students into journal writing and student portfolios. This teacher experienced the portfolio process as an important guide for developing writers; even he, already an accomplished writer, continued to improve. Ultimately, all these forms of portfolios found their way back into the school portfolio.

The teacher team concluded that portfolio development created value for individual teachers, groups of teachers, and the entire school. Individual teachers benefited by developing a voice in the reform process, experiencing an increased capacity for reflection about teaching and learning, and passing their new knowledge along to their students by changing their pedagogical styles and their techniques for student assessment and evaluation. Groups of teachers benefited by becoming aware of other teachers' practices, sharing their knowledge of individual students, and collectively creating new knowledge about their shared school context and reform. Finally, the entire school community benefited as fragmented agendas became more unified, student participation in school portfolio making increased, and grant writing became easier.

The Angelou Toolkit: It's About the Children

In another P & E consultant–school collaboration, the university-based researcher worked primarily with the principal. This consultant already had a working relationship with the principal from an existing reading initiative. When the opportunity to serve as the school's Annenberg P & E occurred, she was a natural choice.

This consultant saw her role as helping the principal develop a systematic plan for the school. She recalled early conversations among initiative organizers and funded school representatives as sometimes heated debates about the purpose of the reform and the localized school-based evaluations. Some people believed the reform – and thus the measurement tools – should focus on student outcome data (e.g., scores on the Texas Assessment of Academic Skills, Stanford Achievement Test, or daily grades). Others argued that the evaluation should focus more on telling the story of how the reform was happening in the schools.

This consultant began by trying to demystify the language of traditional research measurement and evaluation. Her priorities were to help her principal focus the reform and understand how to collect evidence of the work. From these efforts emerged a school plan dubbed the Angelou[6] Educational Toolkit.

The toolkit began as a reflective, qualitative exercise. As this P & E recalled, "It was done as a qualitative study because we wanted collaboration and buy-in from everyone." Large pieces of paper, posted in the school cafeteria, listed probing questions across the top and grade levels down the side. The questions were designed to help participants think through what goals they envisioned for the school,

6. All school names are pseudonyms.

Impact of the Houston Annenberg Challenge

School Development and Teacher Learning

- The Houston Annenberg Challenge invested heavily in teacher and administrator professional development.
- Annenberg schools provided a range of activities to foster teacher learning, including Critical Friends Groups, literature study groups, writing groups, teacher action-research teams, and professional academies.
- Substantial evidence demonstrated that the Houston Annenberg Challenge investment in professional development positively impacted teaching and learning in funded schools.
- Teachers reported that using Annenberg funds to attend conferences helped them to "raise the bar" in their teaching, forcing them to think on a conceptual level and to assist their students in doing the same.
- Teachers collaborated with each other to improve their instructional practices and to create innovative, integrated curriculum lessons.
- Teachers engaged students with active learning strategies by encouraging them to explore and experiment in natural settings.
- Teachers integrated art into the curriculum to engage students in visual, kinesthetic, and auditory ways of learning.
- Teachers used familiar and relevant examples drawn from students' daily lives to connect students to new knowledge.
- Teachers exposed students to complex subject matter and skills while holding them to high achievement standards.
- Teachers helped students to develop critical thinking skills that ultimately deepened students' understanding and expanded their knowledge.

Building Support for Systemic Change

- The Houston Annenberg Challenge provided support to Critical Friends Groups (CFGs) in the metropolitan Houston area by training CFG coaches, principals, and CFG members.
- The Houston Annenberg Challenge implemented a Principals' Leadership Academy to enable area principals to create personal professional development plans and to develop leadership skills.
- The Houston Annenberg Challenge developed new educational programs by creating partnerships such as
  - K–5 Mathematics Specialist Partnership with ExxonMobil and the Houston Independent School District to strengthen teachers' content knowledge and instructional strategies;
  - Partnership for Quality Education with four local universities, six school districts, and one community college to restructure teacher preparation programs;
  - Schools for a New Society with the Carnegie Corporation and the Houston Independent School District to improve the district's twenty-four comprehensive high schools;
  - Passport to Success with the Annenberg Foundation and the United Way of the Texas Gulf Coast to design and implement an after-school program providing an enriched learning environment.

Student Outcomes

- Annenberg-funded schools usually outperformed non-Annenberg-funded schools in the Houston Independent School District.
- Beacon schools led all other schools in reading and mathematics achievement.
- Students in Lamplighter schools exhibited significant positive increases toward mastery.
- Trend data analysis on Beacon and Lamplighter schools indicated that these schools have narrowed, in some cases considerably, the gap between minority and nonminority students in reading and mathematics.
- Beacon and Lamplighter schools also have narrowed, in some cases considerably, the achievement gap between students of different ethnicities and socioeconomic status in reading and mathematics.
- There was little, if any, change, however, in closing the gap between students whose native language is not English and native English speakers.

BEST COPY AVAILABLE

what strategies (i.e., programs or activities) they wanted the school to adopt, and how they would measure progress toward improved outcomes for students and the school. Administrators, teachers, parents, and even some of the fourth- and fifth-graders responded to the questions by posting sticky notes in the appropriate boxes. After two weeks, the P & E removed the charts, typed the responses, and reposted the results. After another round of reactions and responses, the toolkit emerged.

The toolkit became the centerpiece of the school's Annenberg reform work. As a working plan, the toolkit's matrix focused the school's reform work around six interrelated strategies centered on literacy: a schoolwide instructional focus on literacy; resident staff developers; literacy lab; writer's workshop; phonemic awareness, guided reading, and literature circles; and parents as partners.

The toolkit defined each strategy, identified key activities, cited evidence of work, and detailed measurement of progress. For example, the schoolwide instructional focus on literacy was described as "vertical and horizontal teaming that allowed for discussion, planning, and implementation of literacy goals and objectives." The key activities to support this goal (e.g., resident staff developers or literacy lab) made up the substance of the rest of the plan and were intended to build a campus culture to support literacy development. The toolkit offered additional activities, such as principal networking and training, vertical and horizontal teaming across the curriculum, and teachers becoming writers of all genres. Evidence of these activities would be ongoing dissemination of new research to staff, individualized reading with the goal of all students becoming independent readers at grade level and then moving to literature circles, and the presence of writers actively engaged in the writing process. Finally, progress would be measured by professional development documentation, yearly standardized testing, and nonstandardized measures, including Accelerated Reader and literacy lab assessments.

The school developed a toolkit that defined each strategy, identified key activities, cited evidence of work, and detailed measurement of progress.

The school's plan succeeded as demonstrated by higher standardized-test scores, including dramatic gains for previously low-achieving students and an extraordinary reduction in the number of children labeled as learning disabled – from eighty to nine in five years.

As the consultant recalled:

> We were always focused on the children. It's about the children. What do we need to do to help all of them? To meet each of them where they are? We all had this intense support of children. It's about children. It is not about self-aggrandizement.

The P & E recalled how the toolkit matrix led to the idea of resident staff developers:

> We knew we needed more than just literacy. We needed a resident staff developer, and so we conceived this notion of "just-in-time" staff development – [the principal] coined it – because we knew we couldn't do the reform initiative if we waited for whichever day that in-service was scheduled. And we couldn't always justify spending a whole day on whatever it was that the teachers might need. A particular teacher might need help in a content area right now; she might not need it in May when an in-service day is available. So [the principal] used zero-based budgeting to design a way to have a staff developer in each of the content areas.
>
> These resident staff developers met with grade-level teams and content area teams so that the training could be vertically and horizontally

integrated. Sometimes they would meet daily or as needed. They helped to tie everything together and make everyone in the school responsible for outcomes for each and every student. It's everyone's responsibility. It's not just your responsibility because the child happens to be in your classroom right now.

Authentic Assessment: "Deprivatizing" the Work

A third P & E consultant worked with a newly formed experimental high school. Designed around the ten Common Principles advocated by the Coalition of Essential Schools,[7] this small high school focused on developing an environment of personalized learning, using an integrated curriculum, and incorporating service learning for all students. The school's founders created a philosophy for the school based on the idea of authentic assessment rather than traditional graded classes. The faculty initially found this innovative approach very exciting. However, by the second year, many administrators and teachers became overwhelmed as they began to understand the enormous scope of the task they had taken on.

Declaring that she was "single-mindedly focused on the restructuring" at the high school, the P & E was particularly interested in the espoused school goal of creating a professional community that would blur the lines between traditional teacher and student roles. The faculty selected five elements to characterize their vision of professional community: shared norms and values, collective focus on student learning, collaboration, "deprivatized" practice, and reflective dialogue. The P & E began by interviewing the school's sixty faculty and staff members, summarizing the results, and presenting the findings to them as a group. She saw her role as helping the faculty continue to think through the process of putting their vision into practice. She helped the faculty see how their theory of action was evolving, especially in terms of how professional community connected with authentic assessment or espoused theory.

According to her interpretation, the design theory changed each year as new issues arose. The theory in use was best reflected, she said, by the yearly Campus Improvement Plan required by the district. She explained, "I think of theory in use as a way of saying,

The newly formed experimental school's founders had created a philosophy for the school based on the idea of authentic assessment rather than traditional graded classes.

'Did this happen, what happened, what *really* happened, and how do those three things hang together?'"

By the end of the second Annenberg-funded year, the founding principal retired and a new principal was named for this school. The P & E worked closely with the new principal for three years. She described one of her roles as a confidant or "gadfly" for the principal.

> We meet once a month at least and have conversations wherein [the principal] tells me everything that's going on and what she's thinking. And then she asks me to tell her what I think about what she's thinking. I would say I'm her gadfly. Mostly, [she] and I have conversations about her efforts to sort through different ways of encouraging staff to experiment in new ways. Sometimes we talk about how to make them accountable, how she can make them accountable without being overwhelming.

By the time the second principal arrived, a "disconnect" had occurred among the faculty between the philosophical vision for the school and actual practice. Given the extraordinarily complex task of

7. The ten Common Principles are: learning to use one's mind well; less is more and depth over coverage; goals apply to all students; personalization; student-as-worker and teacher-as-coach; demonstration of mastery; tone of decency and trust; commitment to the entire school; resources dedicated to teaching and learning; and democracy and equity. For more information, see <*www.essentialschools.org*>.

What we're doing in this process is deprivatizing; we're forcing people to talk about things that historically don't get talked about on a campus.

creating such a school environment, this disconnect is not surprising. After analysis, the P & E consultant concluded that "clearly defined structures for professional dialogue and continual investigation and analysis" were in place, but the time was not being used productively. Despite the goal of building curriculum and designing instruction and assessment, the faculty did not really seem to know what to do with the designated time. The principal decided to refocus the faculty and staff around the questions, "Who are we, and how good are we at who we are?" These questions proved to be an effective way to launch a collective inquiry process.

As a result of their inquiry, the faculty decided to use Annenberg funding to bring Fred Newmann[8] as a consultant to help them work through their dilemmas with implementing authentic curriculum and instruction. Newmann and his colleague Bruce King from the University of Wisconsin–Madison began three years of consulting by observing staff meetings. The principal remembered Newmann's feedback:

> He and Bruce would sit in meetings, and they would listen. Even though they came to help us understand how to see the world through authentic intellectual work and rubrics, he didn't, because we weren't ready. That fall they showed us disconnects between our goal of authentic instruction and our actual curriculum delivery. Remember our focus on "Who are we?" We believed we were about personalized learning and integrated curriculum. He helped us understand we were not integrating curriculum. Instead, our curriculum was multidisciplinary. Actually, our delivery was developmentally inappropriate because we had ninth through twelfth [grades] combined. If we set the expectations too high – instruction, curriculum, and assessment – then the eleventh- and twelfth-graders could do it, but the ninth- and tenth-graders could not.
>
> Another example was our assessment. We had lots of rubrics, but he showed us our reliability was nonexistent because we never talked about the same rubric. Additionally, our rubrics tended not to tap into substantive, qualitative, authentic work. So what we're doing in this process is deprivatizing; we're forcing people to talk about things that historically don't get talked about on a campus. We were asking people to come forward with lessons they had designed and student work, and we were scrutinizing it. We were actually talking about it. Fred helped us develop a common language and a common understanding about what we were doing and how we could improve.

As Fred Newmann and Bruce King worked with the faculty, the P & E continued her conversations with the principal. The P & E recalled their discussions:

> She's very wonderful to talk to. I mean she's just full of ideas. She's very open, very reflective. Fred – who's a very down-to-earth, accessible person – was stimulating conversations about instruction and learning that hadn't been there before. The principal was trying to rework her role as an administrator by engaging in real shared leadership. And she was concerned about how to encourage instructors to document their instructional practice in such a way that you could tell that changes are taking place. So a lot of our conversation was about that.
>
> For example, during one session I remember she was anxious about evaluating staff. So she talked with me about doing classroom observations. I would ask her questions like, "Is it your intent to have the staff person reflect on what

8 See Newman, Secada, and Wehlage 1995 and Newman and Wehlage 1995.

he/she is doing? Is that your primary motive?" By the end of the conversation, she was seeing things in a broader way. She invited the staff person to come in and give her/his own reflection of the lesson.

All three P &r Es believed strongly in connecting the faculty with current literature. As this consultant explained: "My goal is to think from a theoretical perspective. I try to connect what she [the principal] is saying to literature. She's certainly widely read, but it may be that I have a different sort of theoretical take than she. I try very hard to read enough so that I can bring in something different. My sense is that that enlarges the conversation so she can go with her thoughts to a whole new venue. I think at the end of our conversations she understands better what she is thinking, which is probably the most valuable thing I do there."

WHAT WE LEARNED

We have learned that, for constructive partnerships to work, practitioners must understand that reform is a developmental process. It is a process that takes many turns and leads to unpredictable outcomes. Even though practitioners use guides and specific goals, change is always constant, regardless of activities and strategies. These activities sometimes yielded unexpected positive (or negative) results, and teams learned from their mistakes as well as from their successes. We learned that leaders build feedback loops for constructive criticism and minimize negative personal consequences. Often these strategies resulted in profound shifts in the school organizational culture. When educators experienced these cultural shifts, they were more likely to sustain reform and to engage in continuous improvement activities.

Often these strategies resulted in profound shifts in the school organizational culture. When educators experienced these cultural shifts, they were more likely to sustain reform and to engage in continuous improvement activities.

We have also learned that researchers and practitioners need to understand fully the reform program. They cannot have different sets of operating assumptions. The reform may not work if the researchers and practitioners do not have a similar frame of reference or if the frame of reference is not comprehensive. Both practitioners and researchers need advance training about the nature of the reform, general expectations, and theoretical assumptions. Moreover, they have to agree on a pluralistic methodological orientation – the use of many methods – to yield significant outcomes. Finally, both have to assume that the researchers will provide candid feedback to help with design and implementation of the reform. We believe that these shared assumptions are essential to building successful partnerships.

Finally, we believe that each school needs to have latitude and autonomy in selecting partners. Educators need to trust their schools' partners and to believe they are committed to making serious recommendations for school improvement. Trust and credibility are crucial to the success of the reform effort.

References

Argyris, Chris, and Donald A. Schön. 1992. *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass.

"Child-Centered Schools Initiative." 1996. A Proposal to the Annenberg Foundation. Houston: CCSI.

Craig, C. J. 1997. "Telling Stories: A Way to Access Beginning Teachers' Knowledge," *Teaching Education* 9:1, 61–68.

Craig, C. J. 1999. "Parallel Stories: A Way of Contextualizing Teacher Stories," *Teaching and Teacher Education* 15:4, 397–412.

Cushman, K. 1996. "Documenting Whole-School Change in Essential Schools," *Horace* 12:3 (January), 159–165.

Newmann, F. M., and K. Sconzert. 2000. *Improving Chicago's Schools: School Improvement with External Partners*. Chicago: Consortium on Chicago School Research.

Newmann, F. M., W. G. Secada, and G. G. Wehlage. 1995. *A Guide to Authentic Instruction and Assessment: Vision, Standards and Scoring*. Madison, WI: Wisconsin Center for Education Research.

Newmann, F. M., and G. G. Wehlage. 1995. *Successful School Restructuring: A Report to the Public and Educators by the Center on Organization and Restructuring of Schools*. Madison, WI: Wisconsin Center for Education Research.

Reyes, Pedro, and Joy C. Phillips. 2001. *The Houston Annenberg Challenge Research and Evaluation Study: Year Two: Transforming Public Schools*. Houston: Houston Annenberg Challenge.

Schön, Donald A. 1983. *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.

Schön, Donald A. 1987. *Educating the Reflective Practitioner*. San Francisco: Jossey-Bass.

Schön, Donald A., and Joseph P. McDonald. 1998. *Doing What You Mean to Do in School Reform: Theory of Action in the Annenberg Challenge*. Providence, RI: Brown University, Annenberg Institute for School Reform.

Texas Education Agency. 2003. "A Guide to the Texas Learning Index (TLI)." Available at <*www.tea.state.tx.us/student assessment/resources/guides/tli.html*>

CHAPTER 3:
TRANSFORMING EVENTS:
A LOCAL EDUCATION FUND'S EFFORTS TO PROMOTE LARGE-SCALE URBAN SCHOOL REFORM

Barbara Neufeld, Education Matters, Inc.

Ellen Guiney, The Boston Plan for Excellence in the Public Schools

The Boston Annenberg Challenge

THE BOSTON PLAN for Excellence in the Public Schools (BPE), a local education fund established in 1984, spent its first eleven years awarding grants to Boston teachers and schools with the hope of improving teaching and, thereby, student achievement. Convinced that this strategy could not lead to sustainable schoolwide or districtwide improvement, BPE's Board of Trustees asked the BPE leadership to consider how to better invest in enhancing student achievement in the Boston public schools. With this charge, in 1995 the BPE hired Ellen Guiney as executive director. Guiney worked with Boston's new superintendent of schools, Thomas Payzant, to develop a model of whole-school change intended to have a sustainable impact on schools and on the Boston Public Schools (BPS) as a system.

The BPE committed all the income from its endowment and raised an additional $3 million, dedicating $4.5 million to a four-year reform initiative it called 21st Century Schools. This approach to whole-school improvement focused on instruction. It was designed to be congruent with the superintendent's plan for school reform and began in 1996 in twenty-seven (20 percent) of the district's elementary, middle, and high schools.

Shortly after the BPE began its work, Boston became an Annenberg Challenge site with the BPE designated as a recipient of Annenberg Foundation funds and fiscal agent for the Boston Annenberg Challenge (BAC). Funds were now available for a BAC director, a second cohort of reforming schools, and additional coach support for the 21st Century Schools. With the advent of the BAC, the BPE's 21st Century Schools became known as Cohort I, while schools beginning their work with Annenberg funds were designated as Cohort II. In November 1997, the 21st Century Schools' approach to school improvement was adopted by the Boston School

This paper was written collaboratively by Barbara Neufeld, Senior Research Associate of Education Matters, Inc., who took the lead in outlining the paper, and Ellen Guiney, Executive Director of the Boston Plan for E...ellence, who reviewed and clarified a number of points. The paper was reviewed by Thomas Payzant, Boston's superintendent of schools, before public distribution. The original version of this paper was presented at the 2000 annual meeting of the American Educational Research Association in New Orleans and it was revised in June 2002. Support and data for this paper came from an evaluation grant from the Boston Annenberg Challenge and a four-year, OERI-funded (#R215U7008) qualitative evaluation of the Boston Plan for Excellence.

Committee as the model for school reform in all of the district's schools.[1]

In two short years the BPE, which had comprised three staff people and a controller in June 1996, had transformed itself into an organization responsible for directing school reform in a cohort of Boston schools, in partnership with the BPS but formally outside its jurisdiction. In today's language, the BPE turned itself into an intermediary organization working in a public/private partnership with the BPE.[2] During the summer of 1999, BPE assumed management of the entire BAC.[3] BPE was now an intermediary organization responsible for school reform in half of the district's schools.

Taking on reform in sixty-one schools and trying to support and improve the district had a significant impact on the BPE–BAC. The organization had developed by learning more about teaching and learning and how to improve them; hiring and training people to support the reform effort from inside the BPE–BAC; linking its efforts to those of the BPS and the rest of the BAC; and developing new political skills and relationships with schools, the BPS, and other external forces to meet the challenge of working overtly to change BPS policies and practices from the outside, while also partnering with the school district and schools.

Those researchers who suggest that intermediary organizations can have a powerful impact on school improvement, such as Fullan (2000) and Cohen (2000), point to the knowledge, skill, and resources that such agencies can provide. We have found scant research that addressed the knowledge, skill, and resources that intermediary agencies might need to fulfill such potential. In this chapter, we describe and analyze what has been involved for the BPE–BAC since it took on an explicit partnership role in school reform. The case study explores the changes in the BPE–BAC that accompanied its metamorphosis to an intermediary organization.

We argue that taking on this role fundamentally changed the work of the BPE–BAC. It placed enormous demands on the organization's human and financial resources, as well as its educational knowledge and skills, while simultaneously exposing the organization to the exigencies of the political context the public school system faces. These exigencies led the BPE–BAC to make hard decisions about positioning itself within this political landscape. Finally, enacting the role of an intermediary organization remains extraordinarily difficult and, we argue, depends on the will and capacity of both the local education fund and the local school district. Nonetheless, the role has shown promise in Boston and may be a viable approach to reform in other settings as well.

To develop this argument, we begin by discussing conditions in Boston that seemed essential to the formation of the BPE's role. Second, we describe key organizational and technical demands on the BPE during the start-up phase. Third, we illustrate demands on the BPE–BAC to sustain and deepen the work by

1. Boston's reform model is now called whole-school improvement rather than whole-school change.

2. Gillian Cohen (2000, p. 1) defines intermediary organizations as "independent bodies comprised of multiple stakeholders [designed] to push systems to change both from within and without." McDonald et al. (2000) argues that "intermediary organizations are invested with special resources to support change, and they are presumed free of ordinary interests and ordinary political pressures. They may be designed to be temporary or to stay in business indefinitely. . . . [To do their work] these private interests need to combine their outsider's perspective and clout with insider access." Fullan (2000) also talks about the important role such external organizations can play in school reform. In some cities, a new organization was created to forward the implementation of an Annenberg Challenge grant. This was not the case in Boston, where the BPE began its partnership with the BPS before Boston received an Annenberg Challenge grant, and the BPE was, at the outset, the fiscal agent for the BAC.

3. Despite the BPE's responsibility with respect to the BAC, there remained a distinction between the two organizations. Each maintained its own governing board. When the BPS and the BPE began joint management of the BAC, cochairs Ellen Guiney of the BPE and Timothy Knowles, deputy superintendent for teaching and learning of the BPS, agreed to keep the BAC governing board in order to maintain broad representation for the BAC. It included, for example, the president of the Boston Teachers Union and representatives of corporate, cultural, and foundation communities. The BAC board was phased out when Annenberg funding ended in 2001. Until that time, the BPE referred to itself as the BPE–BAC. We refer to the organization as the BPE–BAC when talking about the work of Cohort I and II schools. We refer to the BPE when we discuss the organization before or after Annenberg funding.

discussing the organizational structures and activities that the BPE–BAC initiated and implemented jointly with the BPS. Fourth, we assess the strengths of the BPE–BAC as an intermediary organization and the challenges the BPE continues to face. In this vein, we consider the prospects for the BPE sustaining its role as an intermediary organization in school reform. We conclude with thoughts about the implications for other agencies that might want to take on similar roles and relationships with their local school districts.

ESSENTIAL CONDITIONS FOR BEGINNING THE BPE'S INTERMEDIARY ROLE

The BPE and BPS collaboration came at a propitious moment in Boston. The superintendent had laid out a comprehensive reform agenda. An unusually supportive mayor and his appointed school board shared that reform agenda and were willing, to a degree seldom found in urban districts, to let the superintendent lead. They offered their support, including a substantial increase in school funds each year. Strong financial support for school reform also came from local private foundations and corporations.

However, it took more than these factors to facilitate the partnership. In particular, the partnership

The Education Matters – Boston Plan for Excellence Partnership

Education Matters, Inc., and the Boston Plan for Excellence have been working together since BPE determined, very early in planning its new role, that it would need an outside evaluator to assess two aspects of its work in an ongoing way: the effect of its work on student, school, and district performance; and the efficacy of its theory and implementation of school reform. With assistance from Dr. Kay Merseth at the Harvard Graduate School of Education, the BPE presented a sketch of its work to potential evaluators. BPE eventually formed an agreement with Education Matters, based on its past work and its willingness to make its qualitative analysis formative as well as summative. (BPE also engaged Policy Studies Associates in Washington, D.C., to conduct a quantitative evaluation of its work.) BPE anticipated it could benefit greatly over time from Education Matters' experience in studying school reform in urban districts across the country. From the outset, Education Matters and BPE collaborated on developing evaluation questions each year that addressed significant areas about which the BPE needed to learn.

At several key points in the last five years, the BPE refined its work based on Education Matters' reports. Two examples are illustrative. First, the BPE on-site coaching model began with the idea that coaches would know what needed to be done from day to day to support schools, that they would be willing to support BPE's reform framework as laid out, and would be able to work more or less independently. These assumptions turned out to be only partially right, and Education Matters' reports led to the creation of weekly professional development for coaches and far greater co-construction of the work. Ultimately, the reports led to BPE completely redesigning its on-site coaching model and increasing its effectiveness.

A second example was equally important. As the case study in this chapter lays out, BPE's outside organization status means that it has very little authority and influence with schools that do not want to undertake reform work. After a four-year struggle with this issue, Education Matters' reports helped both the superintendent and the BPE recognize that they had to address BPE's status directly. The ultimate solution was for BPE to support only those schools willing to work with it and to take on the cutting-edge reform work that the BPE was evolving.

BEST COPY AVAILABLE

was built on a solid foundation that included certain key components. By detailing these essential conditions for beginning the BPE's role as intermediary organization supporting the BPS, we are not arguing that they are the only conditions under which such a partnership could develop. We suggest, however, that conditions will need to exist that establish the knowledge, skill, and trust of the district and the intermediary agency, such as the following:

- The key actors, Ellen Guiney and Thomas Payzant, knew and respected one another. They shared similar views of what it would take to improve schools so that all children achieved high standards, and they had unusual access to researchers and practitioners who were leading thinkers in urban school reform. Both had also engaged in four years of extended discussions as members of the Pew Forum on Education Reform. As a result, Guiney and Payzant shared a vision of whole-school change[4] and a personal working relationship that enabled them to jump-start the joint venture with trust already in place.[5]
- The personality and self-confidence of the superintendent were also crucial. It was necessary for the superintendent to believe that he could maintain his status as the district's leader even as the BPE worked with the first cohort of reforming schools. It was essential that he be able to make two points clear to the local school committee, BPS central office administrators, school-based administrators, teachers, and parents: first, that he had "signed off" on the ideas undergirding the BPE's reform plan, and second, that although the BPE would begin school reform with one set of schools, the district would subsequently support all schools in the same work.

  Despite Payzant's support for the partnership and his belief that it could create healthy tension within the BPS, many in the central office and schools were troubled by the joint venture. Payzant's confidence in the BPE's work and his conviction that the effort was truly collaborative helped him persuade some that the partnership did not represent abdication of authority or responsibility. Still, the BPE did not have the support of all key BPS administrators, and the BPE's role continues to worry many central office administrators.
- The BPE brought considerable strength to the enterprise for three other reasons. First, it had large sums of money to support the effort, having marshaled more than $4 million for initial work in the 21st Century Schools. Then, when the Annenberg Challenge grant of $10 million was awarded in 1996, the BPE–BAC was designated a major recipient of funds because the Annenberg Foundation thought the 21st Century initiative had "great promise." Annenberg stipulated that $4.5 million of the $10 million grant be used to start a second cohort of schools along the same lines. This recognition from a powerful outside foundation greatly strengthened BPE–BAC's influence with the BPS.

4 The theory of action is that improved instruction will improve student achievement, and that the way to improve instruction is to support teachers at the school as they learn in collaboration with one another. Collaborative learning is facilitated by asking teachers and principals to engage in specific activities, called Essentials, which, when undertaken with support, help to change the social structure of each school. The goal is to have people work together to arrive at a common language, common practices, and common goals for their students. The activities help school staff recognize their own learning needs as well as those of students; they also lead to reorganization of time, student groupings, staff, resources, and measurement of progress. Teachers and principals focus, initially, on one content area, most often literacy, and research "best practices" in that area. Their research and adoption of "best practices" involve them in attending to the content to be taught and its alignment with state and district standards. All Boston schools use as a guide the district's Plan for Whole-School Change that describes the Essentials on a one-page chart. The theory asserts that schools cannot undertake this work by themselves. Rather, they need leadership within the school and support from without (coaching, for example) because the work is often threatening and difficult. Virtually all of the additional funding schools receive is for professional development that integrates teachers' learning with their practice, gives participants ongoing feedback, and makes these activities a whole-school, collegial endeavor.

5 Gillian Cohen (2000) suggests that prior working relationships may not be essential for the successful formation and continuation of an educational intermediary organization. However, evidence from school-university partnerships and professional development schools suggests that organizations spend a considerable amount of time learning about each other, working out differences, and developing trust prior to beginning their core work. At the least, Guiney's and Payzant's experience working together enabled them to begin core work expeditiously.

A second strength related to the first is that members of the Boston business community who served on the BPE board of trustees had learned that an investment over the long haul was needed in order to accomplish meaningful changes in the schools and district. The leaders were seasoned actors in the school reform arena and supportive of a reform agenda that would not demand a "quick fix."

Third, The BPE was an established, respected organization with a history in the schools as a result of its previous grants program. It was not a newcomer or upstart in its involvement with schools.

With these conditions in place, Payzant and Guiney still had to negotiate decisions in order for the BPE to begin work. The district required, for example, involvement in approving the schools included in the first cohort, and also assurance that the schools would be told they were still accountable for their BPS responsibilities. Payzant insisted on this latter point, unmistakably emphasizing that schools would receive no special treatment other than BPE support. The BPS also had to observe carefully to see whether the BPE had the capacity to work effectively with a large sample of schools. After all, the BPE had no track record in this regard, and it was a considerable risk on the part of the superintendent to agree to such an experimental partnership. Negotiating these aspects of the partnership took knowledge, skill, and time. Such negotiations continue as the work moves forward.

ORGANIZATIONAL AND TECHNICAL DEMANDS ON THE BPE IN THE START-UP PHASE

To get the 21st Century Schools up and running, BPE had a number of tasks to complete. It had to design the reform activities in which schools would engage. To do this, the BPE conducted separate focus groups with principals, teachers, and parents, and sought advice from national experts. This process convinced the BPE that the missing parts of standards-based reform were unremitting attention to instruction, accompanied by "best-practice"-oriented professional development for teachers and principals that focused on specific student learning goals linked to standards.

As BPE designed the reform, Guiney met frequently with Payzant to ensure that the initiative supported his thinking and experience. The BPE also had to develop an effective and fair selection process and a communication strategy for notifying schools about the grants. Payzant and Guiney had agreed that the BPE would not "cream" the schools with the most potential, nor take on too many difficult schools to give the design a proper test. Finally, once the schools were selected, the BPE needed sufficient human capacity to engage them in initial reform activities and then to design and lead the implementation of further phases. This led the BPE to provide on-site coach support for the work and to hire a director for the initiative.

On-site, whole-school-change coaches had several responsibilities. During the first year they helped schools to assess their instructional needs, select an instructional focus, create structures to engage teachers in working together on instructional issues, and learn how to use both new and existing resources in creative, instructionally focused ways. Coaches were also to instill a belief in teachers and principals that they could turn low-performing students into high achievers by significantly improving their practice. To provide this support, the BPE had to recruit, select, prepare, support, and ultimately, evaluate the coaches.

The BPE also had to employ additional professionals to support implementation and design the next phases of the work. The BPE "borrowed" a Boston principal, Gloria Woods, who knew the reform com-

Coaches were to instill a belief in teachers and principals that they could turn low-performing students into high achievers by significantly improving their practice.

ponents and had the leadership skills to help develop and direct the work.[6] Though hiring a successful BPS principal was crucial to the initiative, it required adjustments for all involved. Woods had to establish new relationships with her former fellow principals and with central office staff. Similarly, she and Guiney, who still had a significant voice in the development of the 21st Century Schools, had to figure out ways of sharing responsibilities. It took time, stamina, and determination to create viable responsibilities and authority within the BPE.

During the first year, staff scrambled to keep ahead of dozens of unanticipated questions from schools. Most questions related either to how schools could use their funds or to whether schools had to comply with information requests from central office if they had already given the BPE similar information. BPE staff could deal with the former, but the latter presented difficulties. The first-year work put enormous demands on the knowledge and skill of the organization and its staff.

ORGANIZATIONAL AND TECHNICAL DEMANDS ON BPE–BAC OF SUSTAINING AND DEEPENING REFORM

Once the partnership was under way, the BPE learned almost immediately that it had to develop its capacity to respond individually to twenty-seven schools that were beginning reform at different stages of readiness. The BPE–BAC had to balance the culture and context of each school with the first-year requirements of its reform agenda, constantly reassess the extent to which its decisions and supports were likely to lead schools toward improving instruction, and determine what resources and tools would help schools implement the intense focus on instruction.

We next describe some demands on the BPE–BAC that were associated with

- working in schools that had varying capacity to undertake reform,
- school- and district-level factors leading to collisions between the BPE–BAC and BPS approaches,
- the end of the cohort structure in June 2001, the advent of a set of schools called Effective Practice schools, and the start of the BPE's role as a research and development arm of the BPS.[7]

In addition, we discuss how the BPE–BAC organized strategies to facilitate school and district reform. In describing these factors, our goal is to heighten understanding of what it meant for the BPE–BAC to work as an intermediary organization and, therefore, what other public education funds might need to ponder if they or their constituencies consider this to be a viable role for them.[8]

Demands on the BPE–BAC That Arose from Work at the Schools

The BPE–BAC experienced many school-related demands. Each required the BPE to develop additional expertise and often to add new staff.

Ongoing Need for Coaches to Ensure Focus on Instruction

BPE had estimated that schools would have new organizational structures and strategies in place by the end of year one. It would then phase out the role of whole-school-change coaches. However, the BPE learned quickly that this phase of the work required sustained attention and funding over considerably more time. Also, as schools chose their instructional focus, they needed another kind of coach to help teachers implement the instructional strategies they would learn in professional development. Thus, with BAC funds, the BPE created the position of content

6. Woods's school was one of the 21st Century Schools, and she had stood out among the principals by her grasp of how this reform differed from what schools had done before. In retrospect, Guiney realizes that the BPE gained enormous credibility for its work as a result of choosing Woods to direct it. She was not fully aware of the import of this choice at the outset.

7. The end of the 2000–2001 school year marked the end of the Boston Annenberg Challenge as a formal organization in Boston. The BPE–BAC ended and the BPE became the primary intermediary organization once again.

8. This chapter focuses on the kinds of work the BPE–BAC took on as an intermediary organization and the impact on the knowledge and skill needs of the organization. Analyses of the implementation and impact of the different aspects of reform can be found in external evaluation reports prepared for the BPE and the BAC. See Neufeld 1999 and Neufeld and Woodworth 1998, 1999.

coach – a coach with expertise in the school's instructional focus.

For both sets of coaches, the BPE had to figure out

- what coaches needed to know and know how to do;
- how to develop that knowledge;
- how to build on what the coaches were learning through their work;
- how to provide the coaches with helpful feedback;
- how to evaluate their work.

This was a tall order fraught with challenges for an organization that a short time ago had been awarding grants to individual teachers and schools.[9]

BPE required schools to direct their energy toward instruction, but few schools had the knowledge or capacity to do so.

Need for Tools to Implement Reform

BPE required schools to direct their energy toward instruction, but few schools had the knowledge or capacity to do so. As a result, it fell to the BPE to figure out how to help them. For example, if the BPE wanted schools to reassess their use of flexible dollars, it had to devise a way for them to do this. If the BPE wanted schools to use performance data from standardized tests to inform instructional decisions, it had to figure out how to get productive and efficient results.

Most often, BPE responded by creating or finding tools to help schools deal with these problems:

- It developed and tested a Resource Review Guide.
- It developed FAST Track, a computer program that expedited data input and analysis.
- It gave coaches and schools a protocol for looking at student work (LASW).
- It designed a Phase Chart that described developmental stages of implementation so that schools could assess their own progress on the reform Essentials.

With a Comprehensive School Reform (CSR) grant from the federal government, the BPE garnered additional resources to assess and improve the work. The grant also enabled the BPE to create a new tool, a "coaches' binder" that assembled in one place all the information schools needed to connect coaching to whole-school improvement.

Transferring Capacity to School-Based Leaders

Initially, the BPE provided schools with externally funded supports, such as the coaches, that enabled them to focus on instruction. As the end of the grant cycle grew near, the BPE–BAC turned its attention to what schools would need to sustain the new instructionally focused practices without the coaches, for example, and/or how to work with schools to find ways to continue external help when it was needed. Toward this end, it created Instructional Leadership Team (ILT) networking meetings and asked coaches to transfer leadership skills to others in the school. With respect to LASW, the BPE–BAC created a series of professional development sessions designed to help teachers and others learn how to facilitate these sessions.

Facing Weaknesses in School-Based Leadership

Throughout the reform effort, the BPE–BAC faced the reality that a few schools lacked leadership from the principal for implementing reform and/or commitment from the teachers to do the work. In this area, BPE–BAC felt the limits of its authority as an intermediary organization. It could not insist on cooperation. Nor could it remove a principal who appeared incompetent or unwilling any more than it could replace teachers who refused to implement practices voted on

9. Neufeld and Woodworth (1997) elaborated the complexities that arose from the early work in schools and from getting the coaching work right and continued their analysis of the coach role in two later reports (1998, 1999). Neufeld (1999) focused on whole-school-change coaches.

In becoming a school reform intermediary organization, the BPE–BAC had to strive to become the kind of responsive learning organization it hoped the schools would become.

by their colleagues. Given its limited options, the BPE–BAC stopped funding schools that did not demonstrate a genuine commitment to the reform agenda. Assuredly, other schools noted this action. Nonetheless, the strategy of cutting off uncooperative schools did little to improve implementation in others.

Ultimately, for BPE–BAC to remain effective, it became critical, especially during the last two years, for BPS to join it in establishing procedures for applying district pressure and support to schools that seemed uncommitted to or unable to implement the reform. These procedures gave the BPE–BAC "authority by association." When schools did not implement the reform effectively, the BPE–BAC cohort director and the BPS superintendent and deputy superintendents discussed each situation, then visited schools together. This collaboration enhanced the BPE–BAC's authority and the resources brought to bear on schools that were struggling with the reform initiative.

Even with "authority by association," however, the BPE–BAC had to convince schools that its work had merit. To the extent that the BPS, teachers, and principals saw value in what the BPE–BAC proposed, they continued to grant it authority. This meant that the BPE–BAC had to be cautious in developing its tools and strategies and had to figure out the right entry points for reform so that its work would be seen as beneficial. The BPE–BAC had to be aware of the latest research findings and strategies for technical assistance. It had to ensure that each school with which it worked had the opportunity and requisite supports to succeed in reform. In becoming a school reform intermediary organization, the BPE–BAC had to strive to become the kind of responsive learning organization it hoped the schools would become. All of this meant that the BPE–BAC had to take a far longer view of its role in reform than it envisioned at the outset.

Demands on the BPE–BAC That Arose from the District Context

One reason for creating intermediary organizations is to bring to bear on school reform the knowledge, skill, and independence of an organization unhampered by the traditions, policies, and practices of the district or school. However, even though the BPE–BAC was formally independent of the district, its work was influenced and constrained by many of the same factors that impinge on the district and schools. An outside perspective may enable an intermediary organization to highlight policies and practices that thwart improvement, but an outside perspective will not enable the intermediary organization to disregard the district context.

In this section, we detail several examples of how the district context was of great concern to the BPE–BAC and how the BPE–BAC addressed these challenges. In describing the BPE–BAC's responses, we aim to highlight the range of knowledge and skill needed by the BPE–BAC, and we emphasize the ways in which dealing directly with the district context led the BPE–BAC to take a more public and political posture than it had ever previously taken.

Alternate Approaches to Implementing Standards

Both the BPE–BAC and the BPS were committed to implementing standards-based reform in all of the district's schools. But they began their work with different ideas, resources, and capacities for moving ahead.

The BPE–BAC invested much training and time with coaches, teachers, and principals to help them learn how to look at student work in teams, assess the work against standards, and assess the quality of the assignments and strategies teachers had used. In addition, teachers were asked to discuss what follow-up they would use to improve their assignments and instruction to help students improve their work. The BPE–BAC saw LASW sessions as the arenas for discussing issues of the quality of student work and implications for instruction.

In contrast, the BPS had begun to implement standards before the BAC funding began to support the LASW professional development sessions. Some key

central office administrators were not convinced that the BPE–BAC's approach was one they wanted to replicate or could replicate in light of its demands for on-site coaching support. In addition, some approached the implementation of standards-based reform as they would any new policy – through relying upon explicit directions for the use of newly developed tools. For example, the BPS created guidelines for a series of products to be completed by students and Task Descriptions – checklists for teachers to follow as they assigned and graded the students' products. The BPS approach assumed that by providing teachers with explicit directions about the requirements for students' work that met standards it would enable teachers to understand these standards and also move them toward new teaching strategies that would enable students to meet the standards.

These two competing approaches – LASW and products/Task Descriptions – led to problems in Cohort I schools, where teachers objected to having to implement two approaches to standards reform. Given the choice, teachers often preferred the Task Descriptions, which were faster and less threatening to use. The Task Descriptions neither required them to share work in small groups, nor to pay explicit attention to changing instruction. In meetings and private discussions, the BPE–BAC objected to the required products and Task Descriptions, arguing that they were low level, counter to the goals of standards-based reform, and unlikely to lead to changes in teaching and learning.[10] The district acknowledged the limitations of the Task Descriptions and allowed schools to use other rubrics to assess student work. However, teachers continued to use the Task Descriptions.

Approaches to Principal Professional Development

The BPE–BAC and the BPS agreed that principals must be informed about instruction and the new strategies their teachers were learning if they were to have an effective leadership role. Yet, many years into the reform, monthly sessions for all 127 principals were still very brief and lacked a coherent focus, and half the session time was spent on operations issues. As a result, the BPE–BAC held its own principals' meetings. But the demands on principals' time were enormous, and principals urged the BPE–BAC to collaborate with the BPS on their professional development. The BPE–BAC was invited to join the BPS in developing the new approach to professional development for principals and headmasters which began during the 1999–2000 school year. Collaboratively, they planned the principals' summer retreat and all-day meetings held five times during the year. Each meeting included attention to instruction and time for principals to meet in cohort and school-level groups. This represented another way in which the BPE–BAC and the BPS extended their collaboration.

Pressure for Increases in Standardized-Test Scores

In Boston, as in many districts, standardized-test scores have taken on increasingly serious consequences for students and schools. The district has tied students' Stanford Achievement Test, ninth edition (SAT-9) scores to decisions about promotion, retention, and access to specialized remedial services. The state requires students to take the Massachusetts Comprehensive Assessment System tests (MCAS), given in a number of grades and tied to state standards. Beginning in 2003, high school students will have to pass MCAS at the tenth-grade level in order to graduate from high school. As a result, there is enormous pressure on teachers and students to increase students' test scores.

In theory, testing should not conflict with the BPS/BPE–BAC reform strategy, which postulates increased student achievement as the result of good instruction. However, the BPE–BAC faced challenges in sustaining that position and maintaining attention to teacher professional development. First, it could not promise that front-loading attention to teachers' learning would quickly increase test scores. Its approach to reform relied on steady, meaningful improvements in instruction. Second, demands for higher test scores are a reality to principals, teachers, and district leadership. Yet focusing on the immediate goal of improving test scores might divert attention from the long-term reform agenda or even undermine it. Third, the BPE–BAC encouraged teachers to use a range of assessment strategies to determine student achievement and, most important, to provide

10 Based on their work in other districts as well as Boston, Neufeld and Woodworth (1998) specifically addressed the serious limits of Task Descriptions.

themselves with feedback on instructional strategies. These approaches to assessment were more complex than that required by a standardized test.

The immediate problem for the BPE–BAC was how to keep its work from being judged prematurely, on the basis of standardized-test results. It had to resist short-term interventions in favor of helping teachers improve how they teach, especially their lowest-performing students, and implementing performance assessments that would indicate students' progress towards high-quality work. As a result, the BPE–BAC worked to develop and strengthen performance assessments that provide alternate ways of looking at student progress and to develop tools that allow schools to track these data efficiently.

Limitations Related to the District's Negotiated Teacher Contract

The BPE–BAC worked with teachers in the context of their negotiated agreement with the BPS. In implementing its reform agenda, however, the BPE–BAC found a number of contract provisions hindering the learning communities desired in schools. We identified two areas in which the negotiated agreement provided a difficult context. One was the Boston Teachers' Union (BTU) contract provisions for teacher transfer, bumping, and hiring. The contract in force at the time enabled more-senior teachers to appropriate positions from satisfactory new teachers through the transfer process. In addition, there was nothing to require that teachers transferring to a school be selected by the school's teacher selection team or agree to implement the school's instructional focus and practices.

At the very least, these provisions led to two problems. First, new teachers who had worked hard and were making good progress in growth and team involvement may have had to leave the school against the wishes of the principal and other teachers. From the BPE–BAC's perspective, this meant that teachers in whom it had invested large sums of money could be lost to the schools. Second, this contract provision made it difficult to create sustainable teams of teachers to implement new instructional strategies. Having to incorporate a teacher new to the school who either was not yet trained in that school's program or who refused to learn and use the program contradicted basic assumptions undergirding the reform.

Another problem was that the BTU contract stipulated that two-thirds of the teachers in a given school must approve changes such as schedule changes. This policy meant that one-third plus one teacher could stymie a faculty's desire to create a block schedule or to bank time for professional development. In addition, teachers had considerable discretion over how they used their planning time. The two-thirds contract provision led to situations in which well over half of the teachers in a school wanted to make schedule changes to support reform but were stymied by their dissenting colleagues. The BPE–BAC was under considerable internal pressure to figure out how to address these issues. Yet, the BPE–BAC, as an intermediary organization, was not the entity that would negotiate a new contract.

Limited Capacity at the District Level

Although some high-level central office administrators participated in professional development for principals/headmasters, these learning opportunities did not always engender the depth and breadth of understanding they needed to implement the reform agenda. As a result, BPE–BAC found itself somewhat at odds with BPS central office staff in two arenas.

First, some central office administrators who worked directly with schools and/or teacher leaders still viewed the implementation of standards-based reform differently than the BPE–BAC. For example, BPS's design of the standards-facilitator role, in place until the end of the 1998–1999 school year, did not encourage the kinds of teacher reflection about practice that the BPE–BAC deemed essential.[11]

Second, BPS internal policies and practices devolved responsibility for professional development to many different units within the BPS. Until the 1999–2000 school year, the district had not developed an overarching plan for teachers' professional

11. See Huebner and Neufeld 1999 for a detailed discussion of this role and the limits of central office capacity to support it.

development or ensured that teachers' learning opportunities were connected to the district's reform agenda.[12] As a result, professional development did not always support whole-school reform efforts and sometimes even contradicted the kinds of instructional practices schools were trying to implement.[13]

The BPE–BAC could neither operate as if these central office professional development practices did not exist, nor put alternative strategies in place. Instead, the BPE–BAC turned to new organizational structures.

Professional development did not always support whole-school-reform efforts and sometimes even contradicted the kinds of instructional practices schools were trying to implement.

BPE–BAC-Initiated Strategies to Facilitate the Work of School and District Reform

Throughout the development of its partnership with the BPS, the BPE–BAC initiated a number of interorganizational structures that had been part of BPE's original proposal. These structures included two standing committees: the Annenberg Working Group (AWG), made up of the partners participating in the BAC, largely key BPS and BPE–BAC reform leaders; and the Resource Action Team (REACT), made up of two principals, top BPS administrators responsible for funding and personnel, Ellen Guiney, and BPE's policy director. These committees met in addition to regular one-to-one meetings between BPE–BAC and BPS.

The structures had two goals: to increase coherence and communication to the schools about school reform, and to address systemic issues that interfered with schools' efforts to improve instruction and increase student achievement.

Annenberg Working Group

The BAC was a complex arrangement of players from inside and outside the system from the outset. Besides the BPE, the BAC included Harvard University, which received a major grant; the Center for Collaborative Education, which coordinates Boston's Pilot Schools – its in-district, charter-like schools; and the Center for Leadership Development, which is run by the BPS but jointly governed by the Boston Teachers' Union. The AWG was created to bring together these partners to monitor implementation and to keep the reform effort moving forward constructively and consistently. The superintendent agreed to attend at least half of each meeting and, for the most part, did so.

The AWG aimed to get beyond logistics and information sharing to become a problem-solving body. To this end, it can cite two important accomplishments. One was its collaboration on the Plan for Whole-School Change chart, adopted by the Boston School Committee, to unify the reform effort across the district and guide schools in defining what constitutes evidence of an improving school. Second, the group advised the superintendent on important issues such as how to measure a school's progress in the reform, what accountability mechanisms to use, and how to implement the literacy models approved by the district for schools' use.

Ultimately, however, the AWG did not function effectively as a problem-solving group. It lacked specific focus and, as a result, dealt with too many issues. It had become too large to work effectively as a group. It tended to operate in a crisis management mode, which meant that its planned agenda gave way to discussions of immediate concerns. Meetings were frequently rescheduled due to time conflicts and the travel plans of key members. Much of the work began to take place in other, more targeted venues. While there was collaboration, it was not taking place within the AWG.

12. For a thorough discussion of limitations of current professional development planning and spending practices in a number of urban districts, see Miles and Hornbeck 2000.

13. During the 1999–2000 school year, the BPS outlined its approach to professional development to the school committee and took steps to align and increase professional development resources targeted to whole-school reform.

Resource Action Team

The superintendent convened the Resource Action Team (REACT) at BPE–BAC's request to study and overcome barriers to schools' effective use of human, financial, and time resources. The REACT team is made up of top administrators, school principals, and BPE–BAC policy experts. As the 21st Century initiative got underway, the schools found that district policies, practices, and contracts often blocked them from using their resources to most effectively improve instruction. REACT's methodology involves researching school-based problems through case studies and bringing to bear interdepartmental expertise to analyze and solve these problems. The case studies start with a problem in a school and build up to an analysis of the systemic forces that led to the problem. Working with the case studies, REACT's members use their expertise to examine the problems schools face and identify how the district needs to change to address those problems. BPE–BAC does the time-consuming work of interviewing and listening to people in both schools and the district offices, then preparing the analysis that illuminates events. The BPS delegates authority to REACT, since the administrators who participate have the power to make changes in policy and practice.

In the first year, REACT documented a host of school-level problems and made recommendations that resulted in some important school-level solutions. Since then, REACT has concentrated on fixing root problems rather than those of single schools. It has effected several major policy changes. For example, as a result of a REACT report on the fragmented spending of professional development funds, the district tightened its oversight of professional development and established spending guidelines. At the end of school year 2000–2001, the BPS and the BPE agreed that it was important to take another look at the coherence of professional development spending. REACT was scheduled to update the initial analysis for the 2001–2002 school year.

REACT also developed a platform for teacher contract reform strategically linked to the district's reform initiative. REACT documented the urgency of addressing these recurring problems of teacher transfer and the required two-thirds votes (detailed above). Though these contract provisions had been considered intractable, the joint deliberations of the BPS and the BPE in REACT led to a decision that they were too critical to ignore. Unlike the professional development report, however, REACT agreed that the case studies and paper pertaining to teacher contract reform should be published independently by the BPE.

Taking responsibility for this publication was a bold step into dangerous territory for the BPE. The BPE needed to figure out how to publish its findings and recommendations without appearing to engage in teacher bashing. Initial BTU response to the BPE–BAC report, circulated in spring of 2000, was extremely negative and resulted in a work-to-rule job action that continued through that school year and into the next.

Nonetheless, several changes advocated by BPE were incorporated into the final contract, and they have made a difference in building collegial teams at schools. The provision that allowed senior teachers to take the positions of first-year teachers who would have liked to stay was eliminated. Important changes to deadlines increased a school's opportunity to hire new teachers who were a better fit with the school's reform plans. Finally, the percentage of votes needed to make major instructional changes (although not the school schedule) dropped from two-thirds to 55 percent, a more manageable percentage to attain.

These examples from REACT demonstrate the potential for an intermediary organization to have a significant impact on the district and schools, the extensive capacity needed by the intermediary organization to fulfill this potential, and the fragility of the intermediary organization's position, in light of its capacity to create conflict as well as resolve significant problems.

REACT demonstrated the potential for an intermediary organization to have a significant impact on the district and schools.

The End of the Cohort Structure and a New Role for the BPE

Spring 2001 marked an important point in the organization of the BPE–BAC itself and in the partnership between the BPE–BAC and the BPS. In May, Superintendent Payzant notified twenty-six schools that they had achieved the status of Effective Practice (EP) Schools (the creation of the EP category was coupled with the end of the district's cohort organization). Their status was made public at a ceremony held at the Federal Reserve Bank in Boston. These schools – drawn largely from Cohorts I and II but with some from Cohort III[14] – had demonstrated considerable progress in implementing the Essentials and had also begun to show clear, important progress in student achievement.

The announcement coincided with the end of the BAC and, for the BPE, the beginning of a new role and responsibilities with schools and the BPS. In this new role, the BPE works exclusively with EP schools to deepen the schools' accomplishments, to document their success so they can inform the work of other BPE schools, and to collaborate with EP schools on initiatives that complement the Essentials.

At the top of the BPE's agenda with EP schools is piloting a new "lab site" model for coaching called Collaborative Coaching and Learning (CCL), to accelerate and deepen professional development.[15] Other plans include developing a network of EP principals and headmasters to identify the learning that helped them to become effective leaders, determine what these principals need to continue their learning, and use their learning and experiences to inform principal/headmaster preparation in Boston. Although this new work will be with only twenty-six schools, it will develop and pilot ideas for potential use across the entire district.[16]

Conclusion: Demands on the BPE–BAC as an Intermediary Organization

The BPE–BAC became a different organization as a result of leading school reform in two cohorts of schools. It greatly expanded its role and its in-house capacity. As the BPE took on the challenge of working as a research and development arm of the district, it had to retool itself for the work at hand. Throughout, the organization had to maintain its relationships with the schools, the district, and its own staff. None of this work has been easy.

The BPE–BAC created a working relationship with the BPS which enabled it to have a significant impact on practices and policies. In doing so, it became both more influential and more vulnerable.

The BPE–BAC created a working relationship with the BPS which enabled it to have a significant impact on practices and policies. In doing so, it became both more influential and more vulnerable. By standing at the forefront of reform and supporting two cohorts of schools, it gained standing among the schools and with the district. By allying with the district on important issues of policy and practice – such as professional development spending – the BPE–BAC increased its value to the superintendent. By taking a public position on contractual issues – albeit one supported by the BPS and the mayor – it incurred the wrath of the BTU which could stymie the BPE–BAC efforts. The BPE–BAC gained considerable influence, but it had to consider the political nature and consequences of its actions more than ever.

14. There were, in all, four cohorts of reforming schools in Boston. Cohort II began the year after Cohort I, etc.

15. This model is based on work being done in New York's public schools by Lucy McCormick Calkins at Teachers College, Columbia University. Each EP school will have the full benefit of a coach with expertise in aspects of Readers' and Writers' workshops for several six-week cycles during the year. The workshop approach is one of the superintendent's priorities for the 2002–2003 school year.

16. This section was adapted from a memo presented to the board of trustees of the BPE on October 2, 2001.

STRENGTHS AND CHALLENGES: THE BOSTON PLAN FOR EXCELLENCE AS AN INTERMEDIARY ORGANIZATION

The BPE–BAC was able to work effectively because it brought certain strengths in short supply within the BPS and most other urban school districts. We turn next to these strengths and to the challenges the BPE continues to face.

Strengths

Stable financial resources. The BPE's permanent endowment gives it a base of operating support and financial independence. In addition, receiving the large Annenberg grant of $2.5 million in the critical early years of reform enabled the BPE to concentrate on designing and implementing the reform rather than on fund-raising.

An ability to stay focused on instruction. The BPE does not have to contend with the multiple issues faced by districts and individual schools, which often get in the way of sustained attention to improving instruction.

Time for reflection, research, and planning. Although the BPE must resist the temptation to become engaged in so much ongoing work that, like its school and district counterparts, it has little or no time left for the important professional activities of reflection, researching, and planning, the organization most often has been able to include these activities as part of its ongoing work. The BPE sees these activities as essential to developing the next steps of reform, and they are supported by the BPE board.

Ability to network outside of Boston. The BPE leaders learned from others engaged in urban school reform through participation in, for example, the Pew Forum, cross-site Annenberg Challenge meetings, and the Public Education Network. Early in the reform, BPE–BAC benefited greatly from a direct connection with strong leadership in New York City Community School District 2 and participating in the Institute for Learning at the University of Pittsburgh. More recently, BPE benefited from a direct connection with Lucy McCormick Calkins, whose work helped spur the development of the CCL model of coaching for the EP schools. These organizations and activities engage BPE leadership in conversations about their work and afford them opportunities to learn from others engaged in similar enterprises.

Freedom to choose which issues to study. Because of its external status and its partnership with the BPS, the BPE has the capacity to identify issues that it and the district deem problematic and then propose ways to address them collaboratively, with the BPE taking the lead in staffing. In this manner, the BPE–BAC initiated the study of professional development discussed above and supported the study of the impact of BTU practices associated with hiring, transfer, and bumping policies. Operations are easier at the BPE; it can make decisions and expend funds more quickly than the BPS. In addition, the BPE does not have to respond to every issue that presents itself, a luxury that school districts lack.

Ability to experiment with new roles and practices. The BPE has the capacity to try out new job possibilities with people it hires and then keep or end these jobs depending on their effectiveness. For example, the BPE–BAC tried to create an administrative assistant position at the high school level to free principals from some administrative work so that they could spend more time on instruction. The position did not work well, and the BPE–BAC terminated it after one year. The BPE–BAC experimented with a range of similar strategies to support literacy development at the high schools. As an intermediary organization, it does not have to commit to a large-scale effort before knowing whether that effort is likely to be effective. It can, as a result, serve as a research and development organization for the district.

The capacity to produce communications for professional development. The BPE designs and produces *Focus*, a newsletter that addresses the progress of reform in Boston and includes informative articles about reform in other districts using related strategies. For example, one issue reported on important research in Chicago about links between the quality of teachers' assignments and the quality of student work. Another issue reviewed the whole-school improvement agenda

and its components to help teachers and principals remember how the parts are connected and are designed to lead to increased student achievement.

These strengths contribute enormously to the BPE's effectiveness, as does the fact that the BPE does not have to run the BPS on a daily basis. The organization has the luxury of being able to focus on a few issues and deal with them in depth. This could not have happened without the agreement and support of the BPS, which granted access to the schools. Nor could the BPE be as strong as it is without the support of its board of trustees, which provides needed flexibility and a reasonable time frame for the work. For example, the BPE board committed the interest on its endowment to support whole-school change and raised funds to support the school-based work. BPE board members take the case for supporting the BPS and its progress to their constituencies. They have been willing to publicize school district problems, such as dysfunctional high schools and a rigid teachers' contract. Board members visit schools at least once each year, developing first-hand knowledge of the work of reform. The board's participation provides the superintendent with a public push to take positions that are unpopular with, for example, the union or the high schools.

Challenges

The BPE has challenges as well as strengths. Some are associated with the organization's place outside of the formal school system, transforming itself as it tries to change schools. Others arise from the exigencies of doing school reform – from the endless issues that need attention and the temptation to address them all. And some come from trying to change policies and practices of the district.

Difficulty of getting teachers on board with the reform agenda. Getting teacher buy-in was a challenge from the outset. This was not a simple task for many reasons, including school cultures in which teacher autonomy was sacrosanct and, therefore, in which individual teachers might feel no compulsion to work with one another in small groups, let alone with the BPE–BAC.[17]

Questions about BPE's authority. BPE must continually nurture and sustain enough authority to do its work. Although it has worked with the BPS as a partner in reform, individual principals and many teachers still consider the BPS as the only authority to shape their work. Indeed, the more work that was demanded of principals and teachers, the more BPE–BAC found it necessary to seek explicit BPS support. Individual principals even suggested that work would have a better chance of being done if the superintendent directed them to do it.[18] Questions about BPE authority become more salient as the reform work continues. It is important that the BPE has now redefined its role in the relationship with the BPS as that of a research and development organization.

Perception of partisanship on behalf of BPE. The BPE must walk a fine line to avoid being seen as an organization co-opted by the district. The more the BPE works in collaboration with the BPS, the greater this danger may become. This will be especially likely if the BPE agrees with the BPS on issues that are not supported by principals and teachers.

The difficulty of assigning credit and shouldering blame. As an intermediary organization, the BPE has to figure out how to deal with "credit" for the reform. An intermediary organization can only be successful if the district is successful, and the intermediary ultimately benefits if the district garners the public credit. The challenge for the intermediary organization is to figure out how to describe its role for funders and researchers so that partnerships are acknowledged and supported. At some point, the BPE may have to respond to district claims that it did not work appropriately toward fulfilling the reform agenda. It may have to shoulder blame as well as garner credit for its work.

17. The state's (and less often the BPS's) strategy of using average standardized achievement scores to determine school performance, a strategy that masks the impact of individual teachers and the impact of instruction on different racial, ethnic, and SES groups, coupled with the fact that in Boston, as elsewhere, student achievement is not a component of teacher evaluation, provided some teachers with a rationale for concluding that their work was already good enough. The BPS use of disaggregated data will now expose variations in student achievement along a number of dimensions.

18. We are not sure that a request from the superintendent would garner greater compliance with BPE–BAC requests. Certainly, schools have been known to disregard requests from "downtown" with considerable impunity.

BEST COPY AVAILABLE

Internal communication issues. The BPE has had to figure out how to communicate to its own staff, including coaches, its overall plans for reform at the schools and how its work at the district level informs those plans. One challenge was the need to lay out to staff and coaches enough of the reform agenda for them to understand the overall plan, while, at the same time, the overall plan is always under development. The inability to solve this communication issue led, at times, to hard feelings between the BPE–BAC leadership and its coaches and other staff. At times, the coaches thought that the BPE–BAC was losing its independence. Internal communications remain an area for further organizational development.

Issues in relationship with coaches. The BPE continues to struggle to get its relationship with coaches right. Hired for their knowledge and skill, coaches have been essential to implementing reform. However, they did not have a hand in designing the reform agenda. This led to different opinions about how to help schools move forward. In addition, coaches were continually frustrated by demands on their time from both the schools and the BPE. The BPE–BAC worked closely with coaches to address these and other issues and achieved some improvements. However, this aspect of BPE's role remains a challenge.

The stretch in capacity needed for districtwide scale-up. Having influenced the BPS to implement new practices across cohorts, the BPE faces the challenge of districtwide scale-up. The BPE–BAC encouraged Cohort I schools to implement performance-based assessments to collect usable data about students' progress throughout the year. Although schools varied in what they produced, they tended to agree that the activity advanced their ability to improve teaching and learning. Performance-based assessments, they agreed, were also an important way of demonstrating student learning that supplemented standardized-test scores. The BPS required schools to implement performance assessments in reading, writing, and mathematics during 1999–2000. This new policy was based on the BPE–BAC's experience with performance-based assessments, the development of the district's promotion policy and transition services (which demanded that schools measure particular students' progress throughout the year), the district's desire for systemwide measures of student progress without waiting for results from the SAT-9 and MCAS, advocacy from several principals for clear expectations about the use of assessment instruments, and the need for the district to identify assessment instruments to help inform instruction. This scale-up required a great deal of work from the schools very quickly, but was not accompanied by sufficient professional development on how to create and analyze performance data in a timely fashion. The challenge for the BPE and the BPS is to find ways to support scale-up that are likely to lead to high-quality practices.

Having influenced the BPS to implement new practices across cohorts, the BPE faces the challenge of districtwide scale-up.

The danger of becoming overextended. The BPE faces the challenge of setting realistic parameters for its work. The BPE's strong partnership with the BPS helped to greatly expand its work. The most dramatic expansion came when the BPE became the BPE–BAC and doubled the number of schools it worked with. This additional school-based work stretched its human resources thin. Scaling back to the twenty-six EP schools was, in part, an effort to avoid becoming overextended and, therefore, subject to the same kinds of capacity dysfunctions that are endemic in the district and schools.

Factors for Future Success

Working with schools to help them improve teaching and learning has been exciting, challenging, and rewarding for the BPE. Working on these issues in partnership with the BPS has, likewise, been exciting, challenging, and rewarding. The work caused the BPE to seek out knowledge and skill with respect to teaching and learning, school organization and culture, and strategies for change. It led the leadership to experiment with ways to transform best practices in literacy instruction, assessment, collaborative school-based planning, and professional development into usable knowledge in a range of schools. And, it led the BPE to work on issues of great significance to the district and its reform agenda.

The partnership has made progress; the role of the BPE is strong. Yet, at the same time, it is potentially unstable and its continued existence depends on

- the ability of the particular people currently in place to get the structural relationship between the BPS and the BPE right so that the relationship outlives the current leadership;
- student outcomes and test scores that support continuation of the reform strategy;
- continued external funding for significant, high-quality coaching and other essential supports;
- the superintendent's willingness and capacity to support internal BPS change that would transform the BPS into a continuous learning organization that can support school reform and improvement;
- the BTU's agreement to changes in supports for teachers, for example, implementation of CCL; and the BTU leadership's success in establishing the Boston union as a participant in school reform;
- the school committee's willingness to continue to support the partnership and role of the BPE;
- the school committee's willingness to continue to shift and add funds to support the reform agenda;
- the school committee's commitment, when the time comes, to select as the next superintendent someone willing to sustain the current reform and work with the BPE as a partner.

To the extent that an intermediary organization is successful, it becomes susceptible to many of the same interests and political pressures brought to bear on the district.

IMPACT OF AND PROSPECTS FOR THE ROLE OF INTERMEDIARY ORGANIZATIONS

Currently, there is great interest in marshaling the forces of external, intermediary organizations to support the implementation of standards-based reform in the nation's cities. The BPE, a local education fund; the BAC, an externally formed organization; and the two organizations operating as the BPE–BAC all fit the description of such organizations as envisioned in the literature. However, as interesting as the idea may be, the potential for such organizations to operate effectively depends on a constellation of conditions that support the role in both the organization and the school district. This role demands a great deal of new knowledge and skill on the part of the external organization, and such organizations are not free of local interests and political pressures. In fact, we would argue that, to the extent that an intermediary organization is successful, it becomes susceptible to many of the same interests and political pressures brought to bear on the district.

BPE–BAC placed itself in the midst of several perplexing questions: How can urban school systems transform themselves so that all students in all schools achieve at high standards? How can we scale up the best of what we know about improving teaching, learning and school organization so that all schools become high-achieving places of learning?

Fullan (2000) argues that we have been unable to answer such questions because we have had too simple an understanding of what it takes to implement reform at the school level and of how the school context, or what he calls the "surrounding

infrastructure," can influence the course of reform. He suggests that

> schools *need* the outside to get the job done. These external forces, however, do not come in helpful packages; they are an amalgam of complex and uncoordinated phenomena. The work of the school is to figure out how to make its relationship with them [parents and community, technology, corporate connections, government policy, and the wider teaching profession] a productive one. (p. 583)[19]

Fullan further argues that schools that figure out how to take advantage of external forces can use them to select and integrate innovations, work constantly on connectedness, choose staff development carefully, and work on applying what they learn.[20] By doing this, schools can overcome two enemies of reform – overload and extreme fragmentation. But Fullan's focus is on the ways in which schools need to figure out and organize what the outside can provide. It is not on what it would take for the outside to actually be organized to provide the appropriate help, nor on what it would take for schools to figure out how to organize the chaotic array of resources.

We agree that schools and districts need to take advantage of what is available in their environments, but we think it is naive to assume that all external, intermediary organizations necessarily have the capacity to support coherent reform that focuses on instruction. Nor do we think it would necessarily be in their interest to do so. As we have demonstrated, taking on such a role fundamentally changes the intermediary organization. Not all intermediary organizations would see this as a benefit.

The story of the BPE–BAC as an intermediary organization is the story of an external force that organized itself and offered its services to a school district and to individual schools. We think it is instructive to consider what it has taken for one such intermediary organization to marshal its forces to get the work done. By doing so, other such organizations might thoughtfully consider whether and how they might take on similar roles in their own school districts.

19. Fullan cites Bryk et al. (1998) as the source of the bracketed list.

20. David K. Cohen (1982) has also written about the role and impact of external, extragovernmental organizations in support of school reform policies.

References

Bryk, Anthony, Penny Bender Sebring, David Kerbow, and Sharon Rollow. 1998. *Charting Chicago School Reform: Democratic Localism as a Lever for Change*. Boulder, CO: Westview Press.

Cohen, David K. 1982. "Policy and Organization: The Impact of State and Federal Educational Policy on School Governance," *Harvard Educational Review* 52:4.

Cohen, Gillian. 2000. "Lessons from the Annenberg Challenge: Intermediary Organizations as Persuasive Agents of Change." Unpublished masters thesis. John F. Kennedy School of Government, Harvard University.

Fullan, Michael. 2000. "The Three Stories of Education Reform," *Phi Delta Kappan* 81:8, 581–584.

Huebner, T., and B. Neufeld. 1999. "Memo on Standards Facilitators." Cambridge, MA: Education Matters.

McDonald, Joseph, Milbrey McLaughlin, and Tom Corcoran. 2000. "Agents of Reform: Role and Function of Intermediary Organizations in the Annenberg Challenge." Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Miles, Karen Hawley, and Matthew Hornbeck. 2000. *Rethinking District Professional Development Spending to Support a District CSR Strategy: Version 5.1*. Resource Reallocation no. 3. Arlington, VA: New American Schools.

Neufeld, Barbara. 1999. "The Boston Annenberg Challenge: Baseline Evaluation Report." Cambridge, MA: Education Matters.

Neufeld, Barbara, and Katrina Woodworth. 1997. "Focus on Coaches: A Memo from Education Matters, Inc. to the Boston Plan for Excellence." Cambridge, MA: Education Matters.

Neufeld, Barbara, and Katrina Woodworth. 1998. "The boston Plan for Excellence's 21st Century Schools Program: First Year Evaluation Report." Cambridge, MA: Education Matters.

Neufeld, Barbara, and Katrina Woodworth. 1999. "The Boston Plan for Excellence's 21st Century Schools Program: Midyear Evaluation Report: 1998–1999." Cambridge, MA: Education Matters.

CHAPTER 4:
BALANCING SUMMATIVE AND FORMATIVE EVALUATION: NEW PARTNERSHIPS, NEW EVALUATION MODELS

Joy Frechtling, Westat

Donald J. Killeen, Transforming Education Through the Arts Challenge

Transforming Education Through the Arts Challenge

EVALUATION CAN take many forms and be undergirded by a variety of philosophies. Stufflebeam (2001) identifies many forms of evaluation, each with its own purpose and set of underlying values: assessments based on the strict discrepancy model (assessing achievements against a set of predefined goals and objectives), goal-free assessments of the phenomena's consequences (looking at the value of what has emerged rather than concordance between goals and outcomes), empowerment evaluations (aimed at energizing a group or segment of the population), advocacy efforts designed to convince audiences of a point of view. Each form looks at evaluation in a slightly different way, with the evaluator assuming a somewhat different role and posture.

Until recently the form considered to be "mainstream" was the discrepancy model, with the evaluator taking the role of an external, objective scientist. Under this model, the evaluator was typically positioned outside of the project, deliberately maintaining a distance from the project as it evolved. Information from the evaluation was shared sparingly – frequently after a project was completed – to prevent what was considered a confounding of an "evaluation effect" with the "program effect."

Increasingly, however, rigid adherence to this model is lessening. And, while the evaluator is still required to be objective and unbiased, evaluators are assuming a more moderate, some would say "less adversarial," posture. The notion of "critical friend" has spread from the technical assistance world to the world of evaluation, creating a function and a challenge for the evaluator that is in many ways more complex than the role of the scientist.

The purpose of this chapter is to present a description of a particular evaluation in the area of arts education reform, charting its evolution from the traditional, mainstream "evaluator as outsider" model to what we believe is an increasingly valuable form of the craft.

THE TETAC PROJECT

In 1996 the J. Paul Getty Trust and the Annenberg Foundation launched a joint effort to explore the impact of joining arts education reform with overall school reform. The effort brought together two powerful forces in educational change: one aimed at promoting the arts in schools, the other at changing the culture and improving the effectiveness of public education. Together they sought to explore whether and in what ways arts education reform and other reform efforts could change how schools are structured, how stakeholders within and outside of the schools interact, the nature of teaching and learning, and the skills and knowledge that students take away from their educational experiences. This jointly supported, five-year effort came to be called the Transforming Education Through the Arts Challenge, or TETAC.

A Framework for Arts Education

America's discontent with its schools began more than four decades ago, following the Soviet Union's launch of the first artificial satellite, Sputnik, in 1957. Two years later, the National Science Foundation gathered thirty-five scientists, scholars, and educators at Woods Hole for a ten-day discussion on how to improve science education in America's schools. The conference generated a model of curriculum reform grounded in the idea that the key educational task of schools was to give students an understanding of the fundamental structure of a *discipline*, defined as any subject with an organized body of knowledge, specific methods of inquiry, and a community of scholars who generally agreed on the fundamental ideas of this field (Bruner 1960, p. 1).

TETAC brought together two powerful forces in educational change: one aimed at promoting the arts in schools, the other at changing the culture and improving the effectiveness of public education.

This model captured the interest of scholars in other fields, such as Manuel Barkan of the Ohio State University, who in the mid-1960s explored how the discipline approach to curriculum reform might be applied to visual arts education. He considered whether art was a discipline in its own right: Did it have a structure similar to that found in the physical sciences? And if it did, how would students engage in a disciplined inquiry in art? Barkan proposed that the structure of art did exist in three domains: art history, art criticism, and art production. He believed that the domains should be treated as equivalent candidates for curriculum attention in arts education, a radical departure from the studio, that is, art production, approach that existed at that time (see Wygant 1993, p. 156).

The movement that grew up around Barkan's approach to discipline-centered art education gained little ground through the late 1960s and had stalled by the mid-1970s, following Barkan's untimely death.

In the early 1980s, the J. Paul Getty Trust decided to establish a center for education in the arts. The Getty Center looked to the work of Barkan as a base for its approach to arts education. To Barkan's three domains, the Getty Center added a fourth, aesthetics, and called these domains the "four disciplines of art education" (Dobbs 1992). This modified approach was formally named Discipline-Based Art Education (DBAE). Proponents of DBAE felt that it provided a rigorous and thorough understanding of an art form.

Initially the Getty Center's efforts focused on visual arts education, but as the effort advanced through the 1980s and 1990s, other arts forms eventually joined the DBAE approach. The Getty Center's goal was to appeal not only to those students traditionally identified as gifted but to a wide range of thinkers and learners. Due to the Getty's efforts, DBAE had a major impact on arts education in the 1980s and 1990s. It became the base from which the TETAC project grew.

The TETAC Approach to Arts Education Reform

As implemented by the TETAC project, DBAE evolved into an approach now frequently called Comprehensive Arts Education (CAE), which retains the focus on the four disciplines promoted by the Getty Center but attempts to link them with current trends in general curriculum reform in several ways.

First, CAE is seen as a way of supporting integrated instruction; that is, instruction that not only includes the four disciplines of DBAE but integrates teaching and learning across subject areas. To achieve this integration, a theme or enduring idea emerging from the arts or life is used as the core around which to focus, integrate, and deliver instruction, not only in the arts, but also in the other core subject areas.

Second, and closely related, arts instruction is viewed as something that is not limited to the arts specialist but can and should be done by all teachers. Under this model, an arts specialist, where one exists, may play the role of a coach or mentor to the regular classroom teacher and may provide students with certain specialized instruction, but the specialist is not solely responsible for arts instruction and arts literacy.

Third, a major facet of CAE is the use of inquiry-based instructional techniques. Like reform efforts in other subject matters, CAE calls for instruction that encourages students to solve problems, take risks, seek alternative solutions, relate learning to real-life experiences, and utilize collaborative as well as individual learning strategies.

Finally, CAE supports new ways of conducting the art and craft of teaching. Collaboration and reflection among teachers are emphasized. Teachers, like their students, are encouraged to try new things, to work and plan together, and to develop both within-grade and cross-grade approaches to instruction that draw on the mutual strengths and knowledge of the members of the educational team.

Comprehensive Arts Education is seen as a way of supporting integrated instruction; that is, instruction that not only includes the four disciplines of DBAE *but integrates teaching and learning across subject areas.*

TETAC was intended to address two agendas: one educational, one political. At the heart of the TETAC project was the educational agenda, designed to show not only that CAE provides a rich and valuable approach to arts instruction, but that it could be the catalyst for improving instruction in all subjects. From the political point of view, TETAC intended to show that the arts belong in the core curriculum of learning in public education, deserving parity with subjects like math and science, rather than being marginalized as an optional part of the education spectrum.

The TETAC Project Management Structure

The management structure for the TETAC project was not typical of reform efforts. To understand the evolution of the project and the role played by evaluation, it is important to understand how the project was managed and how the management structure functioned.

The TETAC project was led by a consortium of six independent regional organizations, each of which oversaw the work of a cluster of six schools in its geographic region. The six organizations were far from uniform. They differed widely in the background and experience of their directors and in the support structures on which they could call. For example, in Ohio the organization was based at the Ohio State University's Department of Art Education and was managed by its chairperson and one staff member, both university-based personnel. In California, the

Annenberg was interested in outcomes, but also in evaluators' engaging in capacity building with the sites, especially helping participants improve their own evaluation skills.

regional organization was directed by a former teacher and an art historian/museum educator and was based at a county office for education.

Decision making and oversight for the project at the national level was carried out by an eight-member national steering committee, composed of one director from each of the six regional organizations and a teacher and a principal representing the project schools. The National Steering Committee chairmanship rotated, with two directors sharing the chair over a staggered, two-year time period. An executive director (also an arts education expert) provided administrative support and other guidance to the project stakeholders and National Steering Committee members. The steering committee operated by consensus and each member had an equal vote.

Finally, the National Steering Committee appointed several national task forces to support implementation. Each task force had a separate charge and a system of mentors, appointed and monitored by each member organization. The mentors were a critical component of TETAC's approach to professional development. They provided direct support to the schools, serving in the roles of technical assistant and critical friend/coach. Who the mentors were, their backgrounds and qualifications, as well as the specific roles they played, varied from region to region.

The national task forces drew on the expertise of mentors from across the project to address cross-cutting issues related to the curriculum and professional development components of the project. This structure allowed for both integrated program oversight and region-specific service delivery; that is, while sharing national goals and operating procedures, there was room for considerable local tailoring of the project's implementation.

The TETAC Schools

The TETAC project began in 1998 with thirty-six schools affiliated with the six member organizations. One school was later forced to withdraw when its district closed the school down, leaving thirty-five: twenty-eight elementary schools, three middle or elementary-middle schools, and four high schools. The schools varied in size, in socio-economic status, and in achievement level.

Some of the schools had previously focused on the arts, while others had no special content emphasis. At some schools, a portion of the teaching staff had been trained in Discipline-Based Art Education under the old Getty grant program (although the trained staff members may have left before the school joined the TETAC project); at other schools, the staff were new to arts instruction. Some schools were involved in multiple reform efforts or special projects, while, for others, TETAC was the only special effort. There were inner-city schools, with all the problems characteristic of urban environments, and suburban schools that were beginning to face new fiscal challenges.

The TETAC National Funders

While many local funders became project partners in their regions, at the national level TETAC joined the investments and goals of two powerful educational funders, the Annenberg Foundation and the J. Paul Getty Trust. At a macro level the funders' interests converged in a desire to improve education; at a micro level, however, many differences, sometimes conflicting, could be found. For example, the Getty Trust placed its highest priority on assessing the impact of CAE on the teaching and learning of the arts, and considered impacts on other areas to be secondary. The Annenberg Foundation was less interested in the arts by themselves and more concerned

about overall changes in schools, school cultures, and connections between schools and their surrounding communities. The Annenberg Foundation was also interested in the potential link between the arts and whole-school reform.

The Getty was interested in an outcome assessment of what a high-quality arts program could accomplish. Getty grants officers believed that they had a mature program, ready to be assessed in a summative evaluation after their many years of work promoting CAE. Annenberg was interested in outcomes, but also in evaluators' engaging in capacity building with the sites, especially helping participants improve their own evaluation skills. Initially, for the Annenberg Foundation the process of change was as important as the impacts of the program on students.

Each of the funders' interests changed quite dramatically over time. In the second year of TETAC's national evaluation, a new management structure was introduced at the Getty Trust, and its interest in TETAC waned. As more pressure was exerted nationally on the Annenberg Foundation to prove the worth of its investments in public education, pressure increased on their funded projects to show results, especially in traditional student achievement.

The TETAC National Evaluation

Although the evaluation was to be jointly funded and monitored by the Getty and the TETAC National Steering Committee through its Annenberg Challenge funds, the request for proposals for the evaluation was developed by the Getty Trust and emphasized all the values and objectives of the Getty's original work in Discipline-Based Arts Education. First and foremost, the evaluation was structured to be an outcome assessment, although interest was shown in assessing other areas, such as the role played by professional development in school change and the strengthening of the schools' staffs in conducting their own evaluation of their work.

Diverging from its previous investments in evaluation, the Getty Trust made a deliberate choice to seek evaluators from outside the arts world who were familiar with school reform and student assessment. Acknowledging that lack of arts expertise could be a problem, the Getty identified expert consultants to support the evaluation team in this area. After a series of competitions (there were several false starts before the Getty figured out what it wanted in an evaluation design and an evaluator), the evaluation contract was awarded in the second year of the project to Westat, a social science research firm in the Washington, D.C., metropolitan area. The Westat evaluators were expected to be objective and unbiased, looking at whether or not the project had attained its stated goals and objectives. While oversight for the evaluation was to be carried out by both the Getty Trust and the TETAC National Steering Committee, the Getty Trust was the dominant partner from the outset.

The four-year TETAC national evaluation effort, implemented in project years two to five, was designed to address the following areas and questions:

Student learning: What impact did TETAC have on student learning in the arts and in non-arts areas?

School climate and culture: How did TETAC affect the school as a place of learning? To what extent was arts education recognized as a critical part of the instructional program? Was there an integration of instruction across subject areas? Was there an environment of inquiry and active engagement?

Implementing the CAE approach: What progress did the thirty-five schools make in implementing the CAE approach? What factors facilitated or hindered the success of this approach in the project schools?

Collaborations: What kinds of collaborations were established? How did these affect the view of arts education? The instructional climate? The instructional program?

Professional development: What types of professional development were delivered to prepare teachers and others in the CAE approach? Who received the support? What areas or skills were covered? What was the impact of such professional development on classroom instruction?

General school reform. What other school reform initiatives were simultaneously under way in the TETAC schools during the project years? What was the interaction between CAE and the other school reform initiatives? What impact did existing school reform initiatives have on the implementation of arts education reform efforts (and vice versa)? To what extent did CAE serve as an agent for broader school reform?

The evaluation used a mixed-method approach, employing both broad-based and targeted data collection. A special feature of the design was an elementary grades arts assessment, developed by the evaluation team members and their arts consultant, which provided a more through examination of achievement in the arts. The design for the four-year evaluation study is shown in Table 1.

Data-Collection Activity	Project Year 2 April 1, 1998 to June 30, 1998	Project Year 3 July 1, 1998 to June 30, 1999	Project Year 4 July 1, 1999 to June 30, 2000	Project Year 5 July 1, 2000 to July 2001
Documents and materials	X	X	X	X
Curriculum units		X	X	X
Teacher survey	X			X
Mentor survey		X		
Student survey	X			X
Site visits	X		X	
Case studies		X		X
Arts assessments		X	X	X
Individual student-achievement data from non-arts assessments		X	X	X

Table 1. Schedule of data-collection activities

THE PATH TO PARTNERSHIP

Within six months after the evaluation began, the evaluation and the relationships between the evaluators, the project leadership, and the funders began to change dramatically. This section describes these changes and the critical events that led to a far different relationship between the evaluators and the project leadership than either had envisioned.

Event 1
The First Evaluation Report: "There's no 'there' there."

A pivotal event in the evaluation and one that set the stage for future interactions was the first evaluation report. In the second year of the TETAC project, shortly after the initiation of the contract in April 1998, the evaluation team began collecting data on where the schools stood with regard to program implementation. In addition to surveying a sample of teachers and students, the team spent one day in each participating school. Each visitation team included an evaluation team member and a staff member of the regional organization working with the particular school. This team approach was deliberately created to draw on the different strengths and knowledge of the arts and evaluation communities. On reflection, this was an initial step in developing a new kind of evaluation partnership.

The daylong site visits were guided by a project-implementation scale, describing features expected to be found in a fully developed TETAC school. The evaluation team built the scale, drawing on information from the original project proposal, from previous examinations of DBAE, and from limited discussions with the Getty grants officer and arts consultant for the national evaluation. The project leadership was also consulted, but they had difficulty translating their vision for the project into measurable terms.

Not being arts experts or from the arts world, the evaluators had to work hard to develop their own understanding of what the project's goals meant, translating such phrases used by the consortium and funders as "art at the core of the curriculum" and "a policy for the arts" into things that could be observed, measured, or asked about. The resulting scale included twenty-two project features that could

be divided roughly into several categories: infrastructure (a leadership team, a strategic plan, a policy for the arts, principal leadership, etc.); instructional practices (planning time, curriculum units, pedagogy, assessment, etc.); supports for instructional practices (personal, material, and physical); and connections to the community (funding, parental support, networking, etc.).

The site visits revealed that, in contrast to what had been billed as a mature program, many schools in TETAC were just beginning to develop their approaches, while only a relatively low number could be considered to have a well-implemented program. On a scale of 1 to 10, with 10 being a mature program, a few of the thirty-five programs could be classified as being at the 6 to 7 range. Most, however, were rated 3 to 4, and some were still in the planning stage in the second year of the project.

The first year's evaluation report to the project's National Steering Committee members and the funders provided a detailed discussion of the status of the project implementation, data on each of the components of project's design, and descriptions of the status of each school site. In the tradition of the external, objective evaluator, the report stated that

> the schools have taken some important initial steps in developing a comprehensive arts education program, but there is a large distance to cover in the next three years if they are to become demonstration schools. Pieces of the program – a policy, a critical mass of trained staff, written units, and assessments – are being put into place slowly and unevenly . . . [S]chools' understanding of how to integrate [CAE] with overall school reform needs to be enhanced and refined; . . . only in some cases do strategic plans provide the necessary scaffolding. . . . Planning is fragmented. . . . Integrated instructional planning has not yet become part of the culture of most schools. Few schools have a firm vision of how their environments might look if TETAC is successful. (Frechtling 1998, p. 39)

The site visits revealed that, in contrast to what had been billed as a mature program, many schools in TETAC were just beginning to develop their approaches.

Event 2
The Project Leadership's Response: Reflection and reinvention

With a relationship barely established, the findings from this first report could easily have resulted in a battle between the project leadership and evaluators, setting the scene for four years of antagonism. However, the project leadership realized the findings were credible, especially since the data were gathered during the site visits with a staff member of the regional organization and the reports on each school had been reviewed with the organization before going into the evaluation report. Indeed, while the project leadership and funders reacted to the evaluation report in a variety of ways, exhibiting differing degrees of surprise, no one called for the killing of the messenger. In fact, they used the data from the first evaluation as an impetus to reflect on the project and clarify its goals, and they sought the input of the evaluation team during their self-examination and problem solving.

When the first evaluation report was presented, the TETAC project had a national project advisory group of experts in the fields of educational reform and school change. The project leadership brought together the advisory group and national evaluators for a two-day meeting to examine the report and reflect on its implications. The meeting yielded several important insights.

First, the process of reform was far more complex than the TETAC project leadership ever dreamed. And while high expectations were important, a grounding in the realities of the change process was also important. The leadership realized that the initial notion

they had shared with the funders and conveyed to the evaluators – that the project had a fully implemented set of programs ready for summative evaluation – simply was not true.

Second, and more important, the leadership realized that, while they shared an understanding of what they were trying to do and how they were trying to do it on the most general level, they had widely varying perceptions of what implementation meant and what a fully formed program would look like.

In reviewing the first evaluation report, the leadership looked carefully at the TETAC implementation rating scale used during the site visits and realized they had never shared ideas about what they were trying to do at this level of specificity. They also realized that in the year and a half since the project was first funded, they had moved away from some of their original ideas about what they were trying to accomplish and what a fully implemented program would look like, but where this movement had taken them was less clear. Further, just as the funders were of two minds about where arts reform and overall school reform stood on the project's agenda, the leadership realized that they, too, had some conflicting priorities and might have taken somewhat different paths in trying to resolve them.

Inadvertently, then, the evaluation and its first report became a catalyst for intensive reflection about the project on the part of the leadership. Instead of becoming defensive or antagonistic, they used the results and the analysis that lay behind it to stimulate their own thinking and to examine the process under way. At this point the leadership took a giant step from being part of a loose configuration of arts education advocates to becoming a real learning community. And, almost without thinking about the alternatives, the evaluation team was brought along as a community member.

Excerpts from the TETAC Final Project Report

From the Findings Section

Analysis of the data collected during the last four years of the project led the evaluators to conclude that the project showed success for the TETAC strategy as it evolved. The findings indicated that the approach holds many benefits for schools able to overcome challenges inherent in any program designed for nationwide implementation. Evaluators measured success on many levels.

Enriching Student Learning

While not for the timid, the TETAC strategy for school reform and the arts clearly provides a means of enriching student learning and for changing the culture of a school. Overall, school reform and TETAC were not only compatible, in many ways they shared the same properties, practices, and goals. Most important, the evaluation showed that the TETAC strategy could be effective in all types of environments, including the inner city.

Providing Flexibility

CAE provided a model for integrated arts instruction that could be adapted to a wide range of teaching and learning environments. It could be integrated effectively with overall school reform, especially at the elementary and middle school levels. The approach was not easy to implement, being guided by general goals and objectives, rather than detailed and closely prescribed practices. However, this gave the approach the flexibility needed to accommodate the variety of mandates schools face.

Experiencing Many Benefits

Schools that embraced the strategy of TETAC, adopting and adapting practices consistent with their local mandates and requirements, experienced many benefits. The improvements included increased collaboration among teachers; more opportunity for thematic, integrated instruction; new ways of teaching and collaborating with students in the learning process; higher expectations for students; and new attitudes about the arts and their value to the curriculum. The CAE approach is effective in promoting deeper learning in the arts. Further, while the evaluation fails to provide evidence of positive effects on learning in other subjects, the findings strongly suggest that adding the arts and broadening students' learning opportunities does not hurt performance in other areas.

BEST COPY AVAILABLE

Event 3
Joining Skills to Solve Problems:
A lesson in synergy

Starting from the dialogue launched by the first evaluation report, the relationship between the project leadership and the evaluation team changed markedly to become more collaborative. While the project leadership remained advocates for the project and steadfastly partisan with regard to its goals, they knew that they had problems to solve and appreciated the insights provided by the evaluation perspective. The evaluators, in turn, became facilitators of the leadership's reflections, assisting, when invited, in their discussions and reacting to new documents being developed. As the leaders reflected on what they meant by project implementation, the evaluators were part of the conversation, both prompting discussion and probing what was meant when discussions were unclear or participants appeared to talk past each other. As the leadership struggled to put to together a renewed mission statement and to clarify the project's goals and expectations, the evaluators acted as second readers of the documents they developed, raising questions and pushing for clarification.

The benefits from this interaction went both ways. While the project leadership benefited from the evaluators' process of inquiry, as well as their knowledge of whole-school reform (a new area for many involved in arts education reform), the evaluators benefited from the project leadership's deliberations on arts education reform and the somewhat diverse slants taken on exactly what TETAC was expected to accomplish. A clear by-product of these interactions was a shared vocabulary and increased understanding of the parlance and jargon of the two worlds.

As plans were made to provide new supports for the schools to carry out the change process, the evaluators became secondary players, commenting on

From the Lessons Learned Section

The five-year evolution of the TETAC strategy encompasses a story of experimentation, lessons, and work begun but not finished as time and money ran out. By sharing the TETAC story and findings, the National Arts Education Consortium hopes to trigger additional experimentation and research that continues to build on this work.

Success Was Achieved on Many Levels

- The Comprehensive Arts Education approach showed great promise for improving learning in the arts and integrating the arts into the core curriculum.
- The TETAC strategy for school reform and the arts provided effective approaches for changing the culture of a school and improving overall instruction while advancing the status of the arts in the regular curriculum. These approaches – including curriculum, professional development and assessment – are not only compatible with general trends in school reform, but they are flexible and adaptable to all types of school mandates and environments, including the inner city.
- The TETAC Visual Arts Assessment Instrument provides a prototype for developing tests to measure learning in the other art forms. Additionally, the quantitative data bank documenting student learning in the arts gathered from the TETAC Visual Arts Assessment Instrument offers a base for other studies.

The full report is available via the Arts Education Partnership Web site at *<www.aep-arts.org>*.

BEST COPY AVAILABLE

Starting from the dialogue launched by the first evaluation report, the relationship between the project leadership and the evaluation team became markedly more collaborative.

structure, attending critical events such as professional development sessions, and helping to examine their efficacy. To a large extent, the balance between formative and summative evaluation changed, with formative considerations playing a much more dominant role as new strategies for project development were explored.

Interestingly, the leadership drew heavily upon some of the tools and processes initially developed for formative evaluation; for example, the project implementation rating scale that was used to guide the site visits and formulate benchmarks for measuring progress. In addition to promoting discussion among the consortium's leadership about what they expected to see in a fully developed TETAC school, it was adopted and adapted by some of the project leadership for their own visits to the schools and used to make their own, interim evaluations of status and change. Their feedback helped the evaluators refine their own approach and validated the scale's usefulness.

Event 4
Figuring Out What It All Means: Determining the legacy

The final report to the nation on the TETAC project was planned as a joint effort of the project's National Steering Committee and the evaluation team. Just as the work over the last three years of project implementation was based on a collaborative relationship, sharing the findings and implications was approached as a joint venture. Some initial ground rules were established and primary responsibilities laid out; for example, the evaluators wrote the findings section of the report, and their voice led in discussion of the summative outcomes. Program voices predominated in the discussion of the program's history and context and took the lead in suggesting next steps in promoting Comprehensive Arts Education.

REFLECTIONS AND LESSONS LEARNED

While the TETAC project's national evaluation was originally designed to be more traditional in nature (i.e., summative), both the project and the national evaluation benefited from the more balanced summative-formative approach that evolved over the four-year evaluation effort. Based on these benefits, several lessons have been gleaned from the TETAC national evaluation effort that might assist others planning a more balanced approach to evaluation.

Lesson 1:
The Importance of an Outside Voice

An unexpected benefit came by hiring a national evaluation firm from outside the main realm of the project (in this case, arts education). Westat surfaced as the firm able to handle the scale of the project's national evaluation. Westat had a good deal of expertise in general school reform evaluation but little expertise specific to the arts education arena. At first, some of the project leadership questioned the logic of hiring a firm new to arts education and debated whether it would be better to use a firm already familiar with TETAC's philosophical foundation. Ultimately,

the leadership decided to hire Westat, since one of their goals was to better link education in the arts with general school reform. They felt that Westat's expertise in this arena would benefit the project, as the findings from the evaluation could be better groomed to reach a wider audience than the arts education community.

Since the evaluators were new to arts education, they needed to probe more deeply than usual into the philosophical underpinnings of the project and its goals and expectations. This information-gathering process soon uncovered a major problem for the project: the national project leadership, drawn from the six regional partner organizations, did not have a common understanding of the key concepts, terminology, and goals outlined in the original project prospectus to the funders. While the leaders all thought that they were in complete agreement, the evaluators' probing showed this not to be the case. This situation meant that as each organization in the consortium designed its initial implementation strategies for realizing the project's goals in its service regions, it was not necessarily in sync with all the other regional organizations and their efforts, making a common evaluation effort difficult at best.

The outside perspective provided by the evaluators' initial information gathering and report gave the project leadership an invaluable reality check early in the life of the project. Since the leadership had thought that they were talking the same language by using the same terminology, it was highly unlikely that they – or for that matter a firm from inside the arts education world – would have uncovered this major problem until far too late in the project's lifecycle. Having the reflections from an outside voice gave the leadership information early enough to allow them to revisit and clearly define a common understanding for the project and to better synchronize their implementation strategies.

For the evaluators, a positive outgrowth from this situation was that very early on they were seen by the leadership as providing important information and guidance that could assist the project in being more successful. Since the evaluators were in the initial stage of clarifying for themselves the project's key concepts, terminology, and goals, they were better able to reflect them back to the project leadership, using a questioning and reflection format for what they were seeing or not seeing based on the original project prospectus. Ultimately, this nonthreatening approach opened up a dialogue between the evaluators and the project leadership that established a productive, nonadversarial working relationship that lasted for the duration of the four-year evaluation effort. This open dialogue gave the evaluators opportunities to build the capacity of the project leadership in relation to various issues of designing and implementing an effective evaluation strategy. It also gave the project leadership opportunities to broaden the lenses used by the evaluators to assist them in analyzing and interpreting the evaluative data that they had gathered.

Lesson 2:
The Importance of Being Specific

When it was evident that the project's evaluation was shifting from a more traditional method to one attempting to balance two sometimes-conflicting approaches – summative and formative – the next challenge facing the project leadership and the evaluators was where the evaluation effort could (or could not) be shifted without damaging the objective quality of the traditional approach. This was a significant issue for the project, since it was essential that the integrity of the eventual findings be accepted in the research and policy communities. This challenge meant that the design and implementation of the plan that would guide the evaluation effort during the four years, as well as the understandings and working relationship between the project leadership

Both the project and the national evaluation benefited from the more balanced summative-formative approach that evolved over the four-year evaluation effort.

and evaluators, had to be spelled out in detail and negotiated up front.

While the leadership and evaluators left the negotiation process open and flexible enough to allow for continual review of the plan and agreements as the evaluation progressed, having a clear blueprint to guide the evaluation effort was essential. Since the project leadership and evaluators had developed a productive process of communication, as discussed in Lesson 1, the head of the evaluation team was included in the leadership's semiannual meetings and monthly conference calls concerning the progress and issues being faced by the project and its evaluation. Together, they worked out a blueprint for the national evaluation effort that paid special attention to several key matters.

To begin with, they decided the goals and needs of the evaluation and potential strategies for realizing them. Next, the evaluators guided the leadership in understanding where formative evaluation approaches might be used effectively, where it was more logical to use summative approaches, and, most important, where it was essential to use a summative approach to assure the integrity of the project's findings. In addition, the evaluators specified the areas of the evaluation that would allow the leadership to collaborate in the data gathering and analysis, and those areas which needed to be left to the evaluators to ensure that bias did not enter into the process. Finally, the leadership asked that the evaluators add an expert in arts education to the team to guide them in the area of arts education reform and to help construct the evaluation instruments for gathering data.

With an outside voice in evaluation and with specific plans, relationships, and understandings, TETAC moved in the direction of a new model for program evaluation. This model combined aspects of traditional research and evaluation design with the value orientation and design modifications that emerged from the culture of systemic reform.

References

Bruner, Jerome. 1960. *The Process of Education*. Cambridge, MA: Harvard University Press.

Dobbs, Stephen M. 1992. *The DBAE Handbook: An Overview of Discipline-Based Art Education*. Santa Monica, CA: J. Paul Getty Trust.

Frechtling, Joy. 1998. "Semi-Annual Progress Report." Prepared by Westat for the TETAC National Steering Committee (September).

Stufflebeam, D. L. 2001. "Evaluation Models." *New Directions for Evaluation* 89 (Spring), 7–99.

Wygant, Foster. 1993. *School Art in American Culture: 1820–1970*. Cincinnati: Interwood Press.

CHAPTER 5:
ESTABLISHING CONNECTIONS BETWEEN EVALUATION AND PROGRAM IMPROVEMENT

Denise D. Quigley
University of California at Los Angeles

Los Angeles Annenberg Metropolitan Project

SPEARHEADED BY LEADERS of the professional, academic, and business communities, a coalition of educators and civic leaders in Los Angeles County came together to take up the challenge offered by Walter Annenberg. In December 1994 the Annenberg Foundation awarded $53 million to the Los Angeles Annenberg Metropolitan Project (LAAMP) to support school reform across several school districts in Los Angeles County, including the Los Angeles Unified School District. The Annenberg award required that LAAMP raise an additional $53 million in public and private matching funds.

CONTEXT AND HISTORY

LAAMP set out to accelerate school reform by building on existing district, community, and school efforts, including the reform efforts of the Los Angeles Educational Alliance for Restructuring (LEARN). LAAMP articulated a theory of action with seven Action Principles[1] for effective learning environments:

1. stable learning communities of intimate scale;
2. a broad, intellectually challenging curriculum;
3. inclusiveness among parents and stakeholders;
4. decentralization of control of resources and decision making;
5. linking professional development to the creation of stable learning communities;
6. reallocation of professional time in schools and families of schools in ways that make it possible for teachers to engage in ongoing conversations about curriculum, pedagogy, standards, and the students themselves; and
7. public accountability.

LAAMP created support organizations known as School Families, designed to link schools to each other more closely than they were in the existing district structure. School Families, on average, consisted of six elementary schools, one or two middle schools, and one high school in the same feeder pattern and geographic area. In all, LAAMP funded twenty-eight K–12 School Families involving 252 schools and about 200,000 students in fourteen school districts. By strengthening articulation across elementary, middle, and high schools and by serving students K–12, LAAMP's leaders hoped that the results of their efforts would remain part of the schools' operations when the Annenberg funding cycle ended.

1. See Goldschmidt (2002) and Gribbons et al. (1999) for details on these principles.

Believing in public accountability, LAAMP emphasized basing decisions on data, using evaluation results, and assessing the effectiveness of School Families' reform efforts by their ability to improve student achievement. According to Kerchner et al. (2000), LAAMP's emphasis on accountability dramatically changed the context of education reform in Los Angeles.

For a comprehensive evaluation, LAAMP contracted with the Los Angeles Consortium for Evaluation (LACE), a collaborative effort among researchers at the National Center for Research on Evaluation, Standards, and Student Testing at the University of California at Los Angeles; the University of Southern California's Center on Educational Governance; and the Claremont Graduate School University. Over the five years of the contract, LACE engaged in several studies, ranging from large-scale, quantitative inquiries into implementation and impact to more fine-grained case studies of particular LAAMP focus areas.

This chapter describes how an external program evaluation team that headed up one of the several comprehensive evaluation studies worked hand-in-hand with program staff and others seeking to build and improve one program within LAAMP, Parents as Learning Partners (PLP). The evaluation events reflected the LAAMP philosophy of accountability. They were also shaped, in part, by the participation of the Weingart Foundation, which made a major grant to LAAMP for parent involvement. The Weingart Foundation required separate external evaluations of the two initiatives it supported: PLP and Design for Excellence: Linking Teaching and Achievement (DELTA), an initiative in teacher professional development. PLP and DELTA were somewhat autonomous within the LACE evaluation because they had a direct funding stream and clients to whom they reported.

LAAMP's emphasis on accountability dramatically changed the context of education reform in Los Angeles.

This chapter first provides an overview of the PLP program and its external evaluation. Next, it reviews the chronology of events to demonstrate how a dynamic relationship among the program, the funder, and the external evaluation team was coupled with keen insight on the part of the program coordinators and program staff to build successful relationships. It concludes with lessons learned.

OVERVIEW OF THE PROGRAM AND THE EVALUATION

The Parents as Learning Partners Initiative

Building on research (Epstein 1995) that identified six types of parental involvement – parenting, volunteering, learning at home, decision making, communication, and collaborating with the community – the Parents as Learning Partners initiative recognized parents as vital participants in the academic lives of their children. The PLP theory of action held that students are more likely to succeed in school when parents and teachers

- have a common sense of shared responsibilities and expectations for students' learning;
- help each other provide learning experiences;
- establish and maintain two-way communication.

This equal partnership between parents and teachers would be apparent in – and measurable by – changes in parents' and teachers' behaviors, perceptions, and attitudes. In turn, these changes would provide a supportive and consistent learning environment for children. As a result, children would develop better attendance, pay more attention to their studies, and form better homework habits, ultimately leading to improved grades and achievement.

Thus, PLP's expansive vision encompassed changes in schools, teachers, parents, and students. The program sought these changes through three types of activities:

- parent education on how to communicate with teachers, assess student work, and provide academic support at home;
- professional development for teachers on ways to involve parents in the classroom, effective outreach efforts, home-learning activities, and volunteer recruitment strategies;
- school services to increase the frequency and quality of parental involvement, such as installing voice mail for parent-teacher communication and establishing lending libraries to encourage parent-child reading.

Within this framework, each School Family and school established specific goals. Some goals were unique, such as expanding child-care services. Many objectives were similar, such as increasing the number, type, and quality of parent workshops.

The External Evaluation Design

The PLP evaluation was one of several LAAMP studies conducted as part of the comprehensive evaluation through the National Center for Research on Evaluation, Standards, and Student Testing at UCLA. The PLP evaluation was unique in its structure at UCLA because it was completely autonomous from the other LAAMP studies. The PLP evaluation had its own funding stream, was designed and run independently by the senior researcher, had its own reporting structure, reported independently to LAAMP, and was treated as a separate project. The only overlap was that it was reviewed internally by the same UCLA reviewer/editor. This independence was a by-product of the fact that the PLP evaluation had its own funding stream and its own independent senior researcher/principal investigator. As a result, the PLP evaluation took a different, more integrated approach than the other LAAMP studies.

Based on the approach of the senior researcher of the evaluation team and the willingness of the PLP director, the evaluation was integrated early on into program development and implementation. A differentiated set of questions allowed the evaluation team to look at the behaviors and attitudes of teachers and parents as students progressed through school. The PLP evaluation tracked a random sample of parent-child pairs from second to fifth grade in program and comparison schools in two of the LAAMP school districts, the Long Beach Unified School District and the Los Angeles Unified School District. A multiyear perspective was well suited to the theory behind the program, which held that a school-family partnership would help students' progress over more than a single year.

The Parents as Learning Partners initiative recognized parents as vital participants in the academic lives of their children.

In the first year (1997–1998), the evaluation focused on implementation and operation:

- Where have the School Families started out?
- What professional development and parent education have School Families implemented/offered that target parent involvement?

In subsequent years, the evaluation focused on impact:

- What are the effects of PLP on teachers?
- What is the impact of PLP activities on parents?
- What is the effect of PLP on students' behaviors and achievement?
- Which family or school processes that affect achievement are enhanced by the parent involvement program?

Integrating the district and school staff at the design stage established key rapport with the schools, cluster leaders, and districts.

Using a rigorous design and triangulation of data, the analyses were based primarily on the comparison of PLP and non-PLP teachers, parents, or students. Regression analyses were also conducted using controls for teacher, classroom, parent, and student characteristics to investigate the influence of PLP on student achievement.

The care that went into designing the comparisons proved important in evaluating PLP impact. Six PLP schools – two from each School Family – were chosen with the help of the local district and program staff to represent the range of parent involvement efforts in the district and area. A comparison school was matched with each program school based on geographic area, parent involvement efforts, involvement in large-scale urban reform efforts, and student and teacher characteristics known from several extant data sources.[2]

Arrangements for obtaining and analyzing student-achievement data were negotiated separately with each district. In finding matched comparison schools, Long Beach Unified School District offered also to match the PLP and non-PLP schools on the basis of prior student achievement; however, in the Los Angeles Unified School District this was not possible.

In the first year, a parent phone survey was conducted in program and comparison schools, selecting a random sample of parents of students in grades two through five. It was administered in seven languages: English, Spanish, Chinese, Vietnamese, Armenian, Russian, and Khmer. The telephone survey was readministered to the parents of the original random sample of second graders when those students were in third and fifth grade. In addition, teachers completed paper-and-pencil surveys when these students were in third and fifth grades – all ninety-five third-grade teachers in 1998–1999 and sixty-two fifth-grade teachers in 2000–2001.

With the encouragement of the Annenberg and Weingart foundations, the evaluation team believed that the evaluation should measure effects on student achievement. This support from the funders was crucial in legitimizing the evaluation team's access to achievement data and the resources needed to analyze those data.

Communication

Integrating the district and school staff in the very beginning of the evaluation, at the design stage, established key rapport with the schools, cluster leaders, and districts. Continuing to work with the district staff, cluster leaders, and principals, although time-consuming, built a base of understanding and support for the evaluation and the initiative that helped on many levels. This support grew gradually, fed by the positive response of teachers, principals, and midlevel district staff about the PLP program and the increasing awareness of the district and school staff that their efforts to partner with parents to support children's learning at home was not impossible and would help student's academic success. This support was nurtured and assisted by the PLP program staff and the evaluation team. Feedback to participants soon took on a life of its own as a powerful part of the initiative.

The PLP program staff also worked closely with the evaluation team. In fact, expecting to use the evaluation data in planning for each school year, program staff found it hard to tolerate even a two-month delay between data collection and reporting. As a result, preliminary evaluation findings were

2. These data sources included: parent involvement items from the local needs survey for teachers conducted in spring 1997; self-reported teacher demographics and teacher-reported student and classroom information from the same survey; school-level student information from the California Basic Education Data System; and school-level test scores on the district's nationally normed tests in reading and mathematics. Teacher and student demographics were compared for each program school against each potential match school.

often presented informally in regular meetings as staff finalized their plans for the upcoming year. Final reports followed and were presented to school staff, PLP staff, and the funder. These communication structures and relationships set the foundation for work that continues today as a result of PLP efforts.

Besides having the program, the evaluation, and the funder married to each other from the inception of the PLP initiative, the PLP program director further solidified this bond by setting up a Core Implementation Team (CIT). The CIT consisted of the funder, the program director, the external evaluation team, one representative from each funded School Family, and the district facilitator for each of the two districts involved in PLP. This team met monthly during the first year and then every other month, with meeting locations rotating to the home site of each member.

The CIT meetings provided an invaluable forum for everyone to roll up their sleeves, ask hard questions, and make things happen. The meetings were a place where the team worked through implementing every aspect of the program; discussed the key elements of the evaluation; organized the work of the schools and districts; updated each other about key events, challenges, findings, and necessary tasks; and focused everyone's efforts on student outcomes. Discussions ranged from events at the school sites and the inclusion or deletion of questions in the evaluation instruments to budgets and timelines and ways of improving the program based on findings. Many improvements to the program were made as a result of the CIT group looking at research results together and discussing implications.

Summary of Evaluation Findings

Over the five years of the Annenberg Challenge grant and the Weingart funding, PLP made progress at various levels (for more complete details, see Quigley 1999, 2000, 2002, forthcoming):

- Teachers and other school staff raised their level of awareness about the importance of parent involvement and set firmer goals.
- More teachers engaged in professional development targeted toward parent involvement and used voice mail.
- Schools continued to offer parent education, and parent satisfaction with the workshops increased.
- Parents contacted and visited their children's classrooms more often. More parents supported homework and reading regularly with their children at home. Parents reviewed children's homework more often and borrowed more education materials from the schools.
- Third-grade students in PLP schools in the Los Angeles Unified School District (LAUSD) scored 5.72 percentile points higher in reading and 4.81 percentile points higher in language arts on the Stanford Achievement Test, ninth edition (SAT-9), in 1998–1999 than students in non-PLP schools. This difference was found after controlling for student performance in reading or language arts, respectively, in the fall and for student ethnicity, income status, Title I status, Limited English Proficient status, third-grade class size, number of third-grade teachers, teachers' emergency credential status, parent education, and employment status of the household. (No differences were found in math scores.)

PLP appeared to influence LAUSD students' reading and language arts performance on the SAT-9 through quality of communication between teacher and parent; parents feeling comfortable at the school and being present on campus, not necessarily in the

Third-grade students in PLP schools in the LAUSD scored 5.72 percentile points higher in reading and 4.81 percentile points higher in language arts on the SAT-9 than students in non-PLP schools.

classroom; and parents providing reading support and general academic support and structure in the home.

Overall, PLP affected student academic achievement by improving the quality of interaction between parents and children in the home and through the quality (not quantity) of interaction and communication between parents and teachers. Importantly, these qualitative improvements were not captured in monitoring the amount of time a parent spent at the child's school or in the child's classroom (as is normally tracked by evaluations of parent involvement programs), but by measuring parent-child interactions in the home and parent-teacher interaction.

EVOLUTION OF THE INTEGRATED EVALUATION APPROACH

Three elements – a rigorous evaluation approach, school-site and macro-level team structures, and effective, multi-way lines of communication – were all integral to the success of the PLP program and have modeled an effective approach for program evaluation. This dynamic is best understood by looking in depth at a chronology of PLP's program development and evaluation.

A rigorous evaluation approach, school-site and macro-level team structures, and effective, multi-way lines of communication – were all integral to the success of the PLP program.

Year 1 (1997–1998): How to define and reach depth of program

The program and the evaluation got off the ground simultaneously, both beginning with planning. LAAMP and PLP first worked with schools to develop action plans and budgets. The plans varied in size, scope, and clarity. Although the Weingart Foundation was pleased with the plans, foundation representatives suggested at one of the first CIT meetings that the plans should take a longer view, focusing on the goals they intended to reach in five years, not just in the first year. This suggestion prompted the schools to realize they should work more closely together.

Schools, districts, and School Families spent time formulating content for parent and teacher education workshops. Different School Families drew on Joyce Epstein's (1995, 2000) materials and tips, as well as Lee and Marlene Canter's (1991) *Parents On Your Side*,[3] but adapted the content of these workshops to fit their particular needs. Some schools and School Families also contracted with PLP-approved service providers for parent education workshops or outreach for events.

Funding flowed very slowly during the first year. As a result, schools and School Families had a difficult time implementing big-ticket or outsourced items, such as the lending libraries and voice mail systems.

In planning for the evaluation, the Weingart Foundation continually stressed the importance of linking student achievement to each school's or School Family's PLP efforts. The need to measure activities against intended goals was also emphasized in many conversations among CIT members, program staff, and district people.

Setting the stage for data collection, the evaluation team at UCLA selected the program and control schools, chose students at the schools through random selection, then designed and constructed the parent phone survey and interview protocols. Data were collected in school-site interviews in the fall and in the phone survey of parents of second through fifth graders in the spring. Demographic and academic performance data were also requested from the Long Beach and Los Angeles school districts.

At the end of the first school year, the evaluation team reported to the CIT on the question: Where have the School Families started out? Based on data

3 *Parents on Your Side* (Canter & Canter 1991) gives educators the skills to gain the support of parents and to follow a consistent approach to working with parents, particularly those of at-risk students.

collected in interviews with principals and school parent coordinators at the program and comparison sites, the presentation stated that:

- It was a new concept for schools to dedicate formal professional development time to parent involvement.
- Parent involvement was seen as an add-on program and solely part of Title I.
- Few teachers were receiving instruction on how to involve parents in the academic programs of their children, even at schools with more developed parent involvement programs.
- Parent education workshops were focused, not on academics or communicating with your child's teacher/school, but on parenting skills.

This information spurred important conversation and reflection about the direction of the program. The CIT discussed how to build depth into the program and how to measure the achievement of its aims. The team realized that PLP had to reach a critical mass of parents and teachers by focusing on *academics*. Parents as well as teachers needed to be informed in order to have higher expectations for students. This would mean more and better outreach to attract parents into activities that would help them support their children's learning.

The Weingart Foundation emphasized that technical assistance funds were intended to build capacity for staff at all levels. Principals should take on more involvement and leadership in PLP. District administrators needed to be better informed and included in program operations. The funder also reiterated that, rather than replacing existing plans and activities, PLP was intended to strengthen them or to create innovative parent involvement activities consistent with goals formulated by each School Family.

During this meeting and subsequent meetings over the summer, it was decided to take the following steps in Year 2:

- In order to analyze program quality, depth of activities, and how activities linked to PLP goals, PLP would require self-evaluation plans from each School Family.

At the end of the first year, the team realized that PLP had to reach a critical mass of parents and teachers by focusing on academics.

- Implementation plans and budgets would be monitored at the school and Family levels.
- The CIT would develop a parent involvement rubric, adapted from Anne Henderson's work, for self-evaluation in conjunction with school-site visits to assess the depth of programs and activities. Schools would rate themselves as novice, apprentice, proficient, or distinguished. Matrices for each site would show the relationship of PLP to overall School Family plans.
- Teacher training was essential to prepare teachers to work with parents. Using *Parents on Your Side* as a model program, all teachers should be reached.
- PLP decided to send program staff, district PLP administrators, and school principals to the "Action Team" training by Joyce Epstein and to implement more of the National Network of Partnership Schools (NNPS) tools she had developed.
- Each district's PLP administrator would need to work closely with the PLP School Family's steering committee or core team in order to oversee PLP activities and promote their alignment with School Family Learning Plans. This meant attending all school-site meetings and CIT meetings.
- Parent mentors/ambassadors and community representatives needed clarification of their job descriptions, duties, and accountability. The funder also emphasized that PLP support of their time was intended for direct outreach to parents.

The internal evaluation early in the second year found that schools faced varying challenges and needs in implementing core elements of the PLP program.

- Outreach activities needed to be strengthened.
- Training for staff at all levels should build the leadership required to sustain programs.
- Data sharing and the communication of evaluation findings should be more widespread. By the end of the summer, the twelve evaluation sites across the three School Families received school-level reports from the evaluation team, known as the "Notes to Principals" report. These reports compared a school with other schools in the PLP Program and with a set of comparison schools chosen as controls. The report compared the schools with respect to PLP and parent involvement activities, plans for parent training and teacher professional development, and parents' comments on how to improve parent involvement at the school.
- Presentation of evaluation findings to key stakeholders was deemed important for gaining and keeping buy-in as well as increasing the depth of the programs.

Year 2 (1998–1999):
How to engage all the needed players

In the second program year, an internal evaluation took responsibility for focusing on the implementation of the PLP program at all sites. The external evaluation would track impact on parents and teachers and the resulting changes in student behavior and achievement.

In the internal evaluation, PLP program staff conducted daylong site visits, interviewing all stakeholders (including parents, teachers, principals, parent ambassadors, teacher's assistants, coordinators, and Action Teams) in focus groups that lasted from thirty minutes to an hour. In addition, the site-visit team reviewed such documentation as meeting agendas, parent workshop schedules, and sign-in sheets. The visits focused on two questions: What challenges is the school facing in creating school/home partnerships to improve academic achievement? What areas of need are common to all PLP schools, and what actions does the project staff need to take to meet the needs of the schools?

The internal evaluation found that schools faced varying challenges and needs in implementing core elements of the PLP program. Some schools did not have enough space to house a Parent Center to offer a large array of workshops. Others were unable to hire a parent ambassador, leaving teachers to shoulder the responsibility of outreach and PLP implementation.

The most common challenge was in implementing effective school Action Teams. Schools were at different stages, with some having no Action Team at all and one school having a team with consistent membership and regular meetings. It was clear that all Action Teams needed more guidance and support to become a strong, permanent part of their schools. Communication between parents and teachers also needed improvement.

The challenges that surfaced in these site visits were brought to the attention of the school and program staff so that they could brainstorm and find solutions. The site visits provided opportunities to look inside the schools and determine what steps PLP should take. They also illuminated many creative examples of how schools were reaching out to parents and increasing parent-teacher communication. These examples were shared across sites.

The external evaluation began the second year by identifying the cohort that would be followed for the rest of the evaluation. Phone numbers and attendance information for the sample of second-graders randomly selected in the previous year (now third-graders) was updated with the help of office managers at the evaluation sites. In addition, this sample was freshened from the original random selection to maintain an adequate longitudinal sample size. Data collection on these students in 1998–1999 included a parent phone survey, paper-and-pencil surveys of teachers in fall and spring, and overall demographic and academic performance data from the two districts (only academic achievement data on the SAT-9 was available from LAUSD). The teacher survey included questions about discipline and reward strategies; communication with parents concerning homework, discipline, and academics; and an evaluation of the child's motivation, academic performance, discipline, and parental involvement.

Early in 1999, the evaluation team completed and shared with the CIT an "Initial Steps" report, a baseline report organized by School Family that compared the PLP evaluation sites with the non-PLP, comparison sites (Quigley 1999). The report included results from data collected in 1997–1998 from the parent phone survey and fall school-site interviews, and overall demographic and academic performance data.

The "Initial Steps" report revealed that PLP schools had raised their level of consciousness about the importance of parent involvement and had set firmer goals in the initial implementation year than non-PLP schools. The parent education workshops and programs were still "business as usual," however, with few enhancements, and there was little focus on academics. There had been trouble with implementing the voice mail system and very little energy spent on building other communication mechanisms among parents, teachers, and school staff.

Professional development for the PLP teachers still did not focus on the kind of parent involvement that specifically supports academics (but prior to PLP, there had generally been no professional development that targeted any sort of parent engagement). One School Family had developed content for teacher workshops and was planning to mandate them in the second year of the grant. The other two School Families had yet to develop a curriculum and plan for teacher professional development.

Coordinating sessions of parent education had been a challenge for most of the school sites. The evaluation team pointed out the need for schools to monitor and track parental participation and to evaluate their own progress and outreach. Outreach to parents and teachers needed more emphasis.

Basically, the report confirmed that "Not much was there" concerning a program that targeted equal relationships between parents and teachers with a specific focus on children's academic learning. In the CIT meeting, however, the evaluation team's report and presentation encouraged the program and schools to translate awareness of the importance of parent involvement into more concrete behaviors. Increasing parent participation and two-way communication between parents and school staff, especially teachers, would require a more concerted effort by program and school staff. They would have to try to reach all teachers and other staff through professional development or other venues, as well as to increase outreach to parents and families.

These themes spurred conversations at the CIT meeting about how to improve communication among parents, administrators, teachers, and staff. The CIT also discussed several items important to the Weingart Foundation's board: implementation teams, outreach to parents, installing communication equipment, number of parents reached, approaching critical mass, strengthening Action Teams of principals and teachers, using self-evaluation forms based on the research of Anne Henderson and Joyce Epstein, and sustainability of the work supported by the grant. As a result of that discussion, several new steps were decided upon and taken during the rest of the program's second year.

- The program created several new avenues for parents to discuss the academics of their child with the schools:

  Coffee with the Principal: an intimate gathering, held on a regular basis, where questions and open discussion are encouraged.

Coffee and Conversation: opportunities for parents and staff members to gather in an informal atmosphere where relationships can be built and barriers reduced.

Tea for Ten: originally developed for parents who could not attend morning meetings; became a set of gatherings where parents of successful students share their parenting, which are then assembled in a format that can be used with other parents.

- In Year 2, twenty-two schools sent teams composed of an administrator or designee, teachers, and parents to the PLP Action Team training.
- Recognizing the need for more teacher professional development in parental involvement focused on academics, PLP facilitators collaborated with the California State University, the largest teacher-training institution in California, to design a curriculum and assist new teachers in this area.
- LAAMP and PLP required school-level action plans instead of School Family plans. These school plans covered goals and activities for five years as well as an implementation plan for Year 2, budget information, technical assistance plans, and governance or staffing changes.
- The senior researcher of the evaluation team or the School Family facilitator visited the schools to present findings to school administration, parents, and teachers at each site.

In general, teachers had lower expectations and a lower assessment of a child's academic standing than did the parents.

The school-level reports helped principals and School Families to see their own progress and challenges in relation to others. Parents' verbatim comments from the phone survey were presented after editing out parents' names to protect the confidentiality of comments about particular staff members. The survey provided the principals with the voices and concerns of their parents. These presentations and reports proved unexpectedly popular. A few principals called to request presentations, and, over time, more and more parents and school staff attended them.

It took a lot of work on the part of the program staff and schools to push the implementation and development of the program to a new level. At the end of the school year, the CIT made plans for Year 3 (1999–2000) in view of the following challenges.

- The challenge of involving teachers was very evident. The surveys of the external evaluation, the site visits of the internal evaluation, and informal discussion with School Family facilitators all revealed that many teachers did not know how to involve parents. More collaboration and integration between DELTA and PLP were suggested. It was also noted that a large disconnect existed between parents and the teachers and administrators who, in most cases, did not live in the school community. PLP needed to help teachers embrace students' achievements and interact with parents on behalf of students, as well as to encourage and help parents advocate on behalf of their children's education.
- Parents and teachers needed a greater sense of urgency about children's achievement. This was evidenced by the mismatch between teachers' expectations and evaluations of student performance and parents' expectations and assessment of their child's skills. In general, teachers had lower expectations and a lower assessment of a child's academic standing than did the parents.
- There was still a long way to go in achieving depth in the program. Although some schools were "getting it," PLP needed to provide leadership in making parental involvement in academics a higher priority for districts. It was thought that PLP could not create change without some tension.

Evaluation could help by providing "disturbing" data, such as low test scores and the disparity in teacher/parent expectations.

- There was general concern about the legacy of the Annenberg Challenge grant, which was approaching its last year. What behaviors and attitudes would have changed? How well would the program be sustained?
- Program staff criticized the "slowness" in the availability of external evaluation data for use in planning. Although the external evaluation team had worked fast, moving from data collection to findings and presentation in a matter of months, the program staff wanted results more quickly.

Three main themes emerged: involving teachers more fully; continuing to increase the depth of the program for parents, teachers, and students; and thinking more seriously about impact and sustainability of the PLP initiative. Each member of the CIT recognized the need to focus on impact, not just progress.

Year 3 (1999–2000): Recognizing the value in evaluation data

Concerns about the program and its evaluation were looming as the third year of the grant began. With the grant ending, LAAMP was surveying its accomplishments and intending to fund only effective activities. This context placed pressure on PLP. In the absence of SAT-9 data, there was no way to link PLP to student achievement. LAAMP and the PLP program recognized that the program's implementation was weak and uneven in many ways. The PLP evaluation, with its longitudinal design and phone survey of a large parent sample was costly, particularly given that it was not yielding results concerning student achievement. There was much discussion within LAAMP, PLP, UCLA, Weingart, and the CIT about how to continue an external evaluation of PLP.

In the interim, the evaluation team and program staff united more strongly to obtain the third-grade SAT-9 data. They sent joint letters to people in the two school districts, underlining the importance of the data for the program and the evaluation.

While options for restructuring and cutting back the PLP evaluation were being discussed, SAT-9 data came in from LAUSD and the analysis began immediately. The evaluation team analyzed the third-grade data from the teacher survey, the parent phone survey, and the SAT-9, finishing a second-year evaluation report that addressed these questions:

- What are the effects of PLP on teachers' practices?[4]
- What is the impact of PLP activities on parents?[5]
- What are the effects on students' third-grade achievement?

Meanwhile, the PLP program staff were deepening the activities of parents and teachers. PLP funded and built innovative activities such as the following into the School Families:

- The *Mother-Daughter College Awareness project*, involving 300 elementary and middle school girls and their mothers at twelve schools. This project targeted girls with the potential to be the first in their families to graduate from college. The girls worked with their mothers to improve their education and life skills, develop leadership abilities, and set high professional goals. Activities included college visits, goal-setting workshops, enhanced adolescent/parent communication, and career planning.
- The *Model Classroom project*, implemented with teachers at four elementary schools to provide teacher professional development, ideas on how to involve parents in classroom instruction and homework activities, resources on student-led conferencing, and mutual encouragement and support.

4. The report described third-grade classroom characteristics based on the PLP teacher survey administered in fall 1997 and spring 1998 to teachers in the six elementary schools participating in PLP and the six comparison elementary schools. Then it presented information on the professional development environment based on school interviews and the PLP teacher survey. Finally, it described interactions between the third-graders' parents and teachers, using the PLP teacher survey and the PLP parent phone survey, administered to a random sample of 673 parents in twelve schools during spring 1998 (when their children were in second grade) and to the same parents in spring 1999 (when their children were in third grade).
5. These findings are based on the PLP parent phone survey.

BEST COPY AVAILABLE

In many schools, teams found that parents needed training to become strong, informed, and committed leaders.

Two other innovative programs – Action Teams and the Parent Curriculum project – illustrate how programs were created and grew. Action Teams had started in the PLP LAUSD schools, where PLP worked with twenty-five schools to create Action Teams designing and implementing systematic, comprehensive approaches to parental involvement. Using Joyce Epstein's conceptual framework (Epstein 2000, Setisinger 1996), the teams involved parents, teachers, administrators, and community members in developing long-term parent engagement plans, training parents to work with teachers and administrators, and providing technology resources.

The school-site teams were fully established at most of the schools within three years of the beginning of the grant. The teams created one-year "action plans" with a budget and a three-year outline of the school's vision for parent involvement. Each team formed subcommittees by goal. Two of the goals and subcommittees had to be academic (one in literacy and one in math) to parallel the student outcome goals. The other two were nonacademic such as family health or partnership subcommittees.

Action Teams at the school sites integrated parent involvement goals with overall plans for school improvement. The meetings were also a venue for delegating tasks and identifying ways of using talent. Most important, regularly scheduled meetings provided time for working out the implementation of a school's vision and plan with all key stakeholders present. In these meetings, where opinions were freely sought and problems addressed, team members developed trust, respect, and support for each other. Stability of team membership, open communication, and the written plan mitigated tension or fragmentation within a team.

In many schools, teams found that parents needed training to become strong, informed, and committed leaders. Parent leadership training was established that enabled parents to participate in decision making regarding their children's education. Activities were put in place to develop leadership and to examine ways of engaging all stakeholders in planning. Parent leaders worked tirelessly to enable other parents to become confident, respected participants at the discussion table alongside teachers and other education professionals.

The Parent Curriculum project grew out of one Action Team's recognition that teachers were often uncomfortable in discussions with parents. The team decided to alleviate this potential communication gap with an innovative program. First, the team designed a portion of the school's teacher professional development to address strategies for communication with parents. Second, the principal, parent coordinator, and fifty teachers developed a Parent Curriculum for each grade level, based on the California Standards. The teacher first teaches the lesson to parents as if the students were in the classroom, so that the parents hear the lesson from the child's point of view, understand the concepts behind the "interactive homework" that will require them to work with the child, and receive instruction on the homework. Then the teacher teaches the lesson to the children, and the parents and children complete the homework together at home. This Parent Curriculum is now being used for the entire School Family. Not only has it improved parent-teacher communication, it has focused this communication on academic achievement and support.

Toward the middle of the third year, the PLP evaluation report with the SAT-9 data was finished. The findings were first presented to a CIT meeting, where the following findings were considered especially important.

- The most dramatic finding was that, controlling for teacher, student, and parent characteristics, LAUSD third-graders in PLP schools scored 5.72 percentile points higher in reading and 4.81 per-

centile points higher in language arts on the 1998–1999 SAT-9 than non-PLP students. Further analysis showed that performance was significantly higher when parents felt welcome at the school, attended schoolwide events, and believed that teachers kept them informed. Students also had higher achievement when parents supported academics, especially reading, at home. Math scores were not significantly different.

- Twice as many teachers in PLP schools (50 percent) had professional development that focused on parental involvement than the non-PLP teachers (25 percent), although they said they were not receiving enough preparation for engaging all types of parents. About 70 percent of PLP teachers were accepting responsibility for parental involvement.
- Communication between schools and parents was generally one-way, with information going out from teachers to parents, in both PLP and non-PLP schools. However, some PLP teachers (25 percent) reported more frequent communication and some two-way communication.
- Both PLP and non-PLP parents had less involvement when their children were in third grade compared with second grade, and they helped their child with homework less often. Decreasing parent participation as students progress through the grades is a trend well documented in other research; however, the decline in PLP parents' participation was less severe than among non-PLP parents.
- More PLP parents felt that parent workshops and activities were helpful and worthwhile, compared with non-PLP parents.

In comparing schools over the two years, the evaluation found many differences in favor of the PLP schools. For example, PLP parents were more likely to drill students at home, read with the child every day, and have someone help the child daily with reading or homework. In the PLP schools, parents received more newsletters and bulletins. More PLP parents attended trainings, especially on parenting skills. PLP parents were particularly positive about workshops on how to help children with schoolwork, communicate with teachers and staff, and practice reading at home. Teachers at PLP schools had significantly more opportunities for professional development targeted at parent involvement, and they believed more strongly that involving parents in education was worth any effort. PLP administrators had a greater desire to involve teachers in engaging parents and understood parent involvement as a process critically important to a quality education.

These findings added up to the beginnings of a shift toward a more inclusive, academic-focused culture. By involving administrators, teachers, and parents in expanding parental engagement, PLP had created a systemic approach that proved change is possible. However, PLP needed to affect a critical mass of parents, and this was not happening. The evaluation team recommended that PLP "hammer home" this need with principals and district staff, target even more teachers through other structures in the school, and get more parents to participate more regularly.

Overall, the CIT members were happy with PLP's breadth of vision, intensity of focus, and the significant differences found in its home learning practices and student achievement. The CIT recognized that the key now was to institutionalize PLP activities and reallocate funds to make parent involvement a core priority at schools.

Having reviewed these evaluation results, LAAMP found funds to survey the parents and teachers as the PLP cohort moved into fifth grade. The evaluation as originally designed was put back into full swing.

Performance on the SAT-9 was significantly higher when parents felt welcome at the school, attended schoolwide events, and believed that teachers kept them informed.

Year 4 (2000–2001) and Year 5 (2001–2002): Back in the saddle again

The external evaluation team collected data in 2000–2001 on the students who were now in fifth grade. The team reinstated the parent phone survey, the fall and spring teacher surveys, and interviews of the program staff. Extra effort went into finding current phone numbers for the sample. Although many of the study sites had new principals who were not aware of the evaluation, all twelve schools consented to be in the study again. The evaluation team also renewed its emphasis on gaining SAT-9 data from the districts, seeking data from the students' fourth- and fifth-grade years. The districts provided the data in early 2002 and analysis got under way.

In Year 4, the PLP program initiated plans for district scale-up to bring the initiative into more schools in the two participating districts. PLP efforts were started in eight additional elementary schools in the Long Beach district as well as eleven additional elementary schools, one additional middle school, and one additional high school in the Los Angeles district. Both districts committed to sustaining PLP efforts until 2006. Besides growing the program into more schools, PLP focused on building Action Teams at more schools, further developing the content of the Mother-Daughter College Awareness project and finishing the Parent Curriculum project for all grades.

Also important in building a legacy from PLP, Families In Schools (FIS) began operation as a new nonprofit organization with the mission of building partnerships of families, schools, and communities to help students achieve academic success. Annenberg Transition funds, obtained from the submission and approval of a transition plan to the Annenberg Foundation, supported the start-up of FIS, along with other funds from public and private sources. The organization has four goals:

- strengthen the capacity of families to support their children's academic achievement and access to higher education
- engage educators to increase and sustain their capacity to partner with families to improve academic achievement
- leverage resources to develop and sustain programs for parental involvement essential to increasing student achievement
- develop and disseminate "best practices" through evaluation, research, and practical experiences

Established in early 2001, by mid-2002 FIS was serving five school districts, six Head Start programs, twelve Early Education Centers, nine Child Development Centers, sixty-four elementary schools, seven middle schools, eight high schools, and 36,000 families. FIS works on family literacy, parent education, family support and awareness of college readiness, teacher training, and the planning and organizing of family involvement at schools and districts. FIS leverages improvement and support in these areas through research and evaluation on impact and program quality as well as sharing best practices and lessons learned.

Sustainability of the PLP effort was shown by the structures in place at the schools and districts and the continuation of school plans, including Action Teams. School districts are funding most of the costs paid previously by the grant. Staff have stayed in place at the school sites and the district level.

In sum, three main elements – a rigorous evaluation approach and design, school-site and macro-level team structures, and effective, multi-way communication lines – were integral to the success of the PLP program and have modeled an effective approach for program evaluation. The next section outlines overarching lessons learned.

Definitive research designs inform programs about how to improve their depth, refine their processes, and leverage support for sustainability.

LESSONS LEARNED

Over the course of the PLP evaluation, the program staff, the CIT, and the evaluation team found that evaluation is a process of looking backward in order to steer forward more skillfully. Several lessons were learned as the external evaluation team worked closely with the program staff, school and district staff, and the funder.

Communication surrounding improvement should not be "ad hoc" but formalized in specific settings and structures.

Lesson 1

Programs need to understand how they operate and what challenges they face as well as their impact in order to be effective at change and improvement. Definitive research designs inform programs about how to improve their depth, refine their processes, and leverage support for sustainability. Using data from internal and external evaluation can help programs achieve their goals. This use of data is also supported by the final evaluation report on LAAMP (Goldschmidt 2002), which found that the schools and School Families that consistently implemented two of the LAAMP action principles – parent involvement and data use – demonstrated higher student achievement.

Lesson 2

District and principal support are key to effecting school change because they set the direction and tone of schools' priorities. Establishing the belief among district staff and principals that parental involvement is critical to children's learning is the key that allows for choices to be made about time and money that create the willingness and opportunity of schools and families to partner. The buy-in and support of both district staff and principals is one of the necessary steps in building partnerships with parents/families that support and improve a child's academic learning environment.

Lesson 3

Multi-way communication with stakeholders at all levels is key for buy-in, for maintaining focus on goals, and for improving programs. Communication surrounding improvement should not be "ad hoc" but formalized in specific settings and structures. Communication needs to be started early, continued often, and directed at all participants to gain a 360-degree view. Interim findings must be delivered promptly and at all levels to keep people on target. With feedback and communication, there is less slippage in implementation, and program development sustains its momentum.

Lesson 4

Building communication lines and improving the quality of parent-teacher communication are the conduits for increasing student achievement. When parents and teachers communicate specifically about how parents can support classroom learning through learning strategies at home, and when parents engage in learning activities with their child, the child experiences a supportive and consistent learning environment, resulting in improved achievement.

Lesson 5

Joining mixed groups of stakeholders in conversations around the tough issues of implementing a program and around interim evaluation findings expedites the building of a program and allows a close watch on its improvement and impact. This type of collaboration is possible in teams where members talk, listen, and celebrate the process and progress of their mission. These teams should exist both at the macro level and the site level of program operation.

Lesson 6

Funders are looking for educators who respond to their questions about change and accountability. It is the evaluation-oriented educators and program staff who will have the ear of people who fund education. This means that the evaluator(s) should continually look ahead to communicate about possible improvements, and that educators should learn to report student outcome data clearly keyed to their mission – the development of children.

Lesson 7

An infrastructure for evaluation begins with human development – developing a mindset and shared plan among stakeholders in support of evaluation use. It requires formal procedures for involvement of all stakeholders; clear roles and responsibilities; means for collecting, analyzing, sharing, and acting on information; and methods for periodically refining the interaction between the evaluation and the program. The will to evaluate in an open, participatory manner is key.

Lesson 8

Evaluators should play an integrated role at stages much earlier than the writing of a final report. Evaluation data can help program development and program improvement by informing program participants and decision-makers. This should act as a wake-up call to evaluators about the scope and type of function they play in improving programs. Interactive, early communication between program and evaluation staff is essential in building an effective program.

In sum, evaluation is a means for strengthening programs, schools, and education. If evaluation is seen as a continual learning process, it becomes an integral part of the renewal intended to make schools more effective and to move them toward their goals and ideals. Evaluation, the systematic judgment of the value of programs, projects, personnel, and other parts of the educational system, is essential. In the words of Michael Scriven (1991, p. 4), "Without such a process, there is no way to distinguish the worthwhile from the worthless."

References

Canter, Lee, and Marlene Canter. 1991. *Parents on Your Side*. Santa Monica, CA: L. Canter.

Epstein, Joyce L. 1995. "School/Family/Community Partnerships: Caring for the Children We Share," *Phi Delta Kappan* 76 (May), 701–712.

Epstein, Joyce L. 2000. "How to Organize Successful Action Teams," *Type 2*, Newsletter of the National Network of Partnership Schools (Fall), 2.

Goldschmidt, Paul. 2002. *A Comparison of Student Achievement in LAAMP and Non-LAAMP Schools in Los Angeles County: Longitudinal Analysis Results, 1997/1998–1999/2000*. CSE Technical Report. Los Angeles: UCLA, Center for the Study of Evaluation.

Gribbons, B., E. Baker, J. Herman, K. Kim, X. Ai, H. Masunaga, and J. Lihosit. 1999. *Monitoring Changes in Los Angeles Schools: Guidance on Indicators*. CSE Technical Report. Los Angeles: UCLA, Center for the Study of Evaluation.

Kerchner, Charles Taylor, Jason Abbott, DeLacey Gamley, and David Menefee-Libey. 2000. *The Impact of the Los Angeles Annenberg Metropolitan Project on Public School Reform: An Analysis of LAAMP Board and Staff Activities*. Los Angeles: Claremont Graduate School University, Los Angeles Compact on Evaluation.

Quigley, Denise D. 1999. *Parents as Learning Partners (PLP) Initial Steps: Implementation and Baseline Findings*. CSE Technical Report. Los Angeles: UCLA, Center for the Study of Evaluation.

Quigley, Denise D. 2000. *Parents and Teachers Working Together to Support Third-Grade Achievement: Parents as Learning Partners (PLP) Findings*. CSE Technical Report. Los Angeles: UCLA, Center for the Study of Evaluation.

Quigley, Denise D. 2002. *Parents Involved in the Academic Learning of Their Children: A Longitudinal Study of Parent-Child Pairs from Second through Fifth Grade*. CSE Technical Report. Los Angeles: UCLA, Center for the Study of Evaluation.

Quigley, Denise D. Forthcoming. "Promoting Third-Grade Language Achievement in an Urban Setting through Parent-Teacher Communication and a Supportive Home Learning Environment," *Journal of Education for Students at Risk*.

Scriven, Michael. 1991. *Evaluation Thesaurus*. 4th ed. Thousand Oaks, CA: Sage.

Setsinger, R. M. J. 1996. "Nurturing Partnerships between Schools and Families." In *Boundary Crossings: Educational Partnerships and School Leadership*, edited by P. A. Cordiero. San Francisco: Jossey-Bass.

CHAPTER 6:
THE PROMISE AND CHALLENGE OF EVALUATING SYSTEMIC REFORM IN AN URBAN DISTRICT

Thomas Corcoran
Consortium for Policy Research in Education

Ellen Foley
Annenberg Institute for School Reform

Philadelphia's Children Achieving Challenge

IN PHILADELPHIA, the Annenberg Challenge initiative, known as Children Achieving, set the ambitious goal of having every student achieve proficiency in three core subject areas – math, reading, and science – by 2008. With the support of $50 million from the Annenberg Foundation and a $100-million double match provided by Philadelphia businesses and local foundations, the School District of Philadelphia set out to achieve this goal by designing and implementing ambitious reforms in almost all aspects of its work, and, in the words of its fervent superintendent, David Hornbeck, to do it "all at once."

In this chapter, we describe what happened in Philadelphia in the late 1990s and what we learned about school reform during five years of studying the Children Achieving initiative. Drawing on both quantitative and qualitative data collected during the evaluation of Children Achieving, we examine its theory of action, its implementation, its successes and disappointments.[1] We also describe our experience as evaluators, observing the initiative at close range and providing formative feedback. We argue that the initiative demonstrated some promising early gains in achievement. However, serious flaws in design and execution and inadequate attention to the Philadelphia context ultimately limited its impact and brought it to an end. In particular, we conclude that the policy dictum that everything had to be done at once, as well as poor sequencing of actions, failure to win teacher support for the reforms, and the emphasis on raising standardized-test scores led to uneven, often superficial, implementation. And we conclude that these flaws affected the evaluation by curtailing opportunities to provide candid, timely recommendations for midcourse corrections.

1. In 1996, the Children Achieving Challenge commissioned the Consortium for Policy Research in Education (CPRE) and Research for Action (RFA) to conduct a four-year evaluation of Philadelphia's Children Achieving initiative. CPRE conducted two systemwide surveys of teachers about the impact of the reforms on their daily work and about the character of their instruction. CPRE and RFA staff members also collected data from forty-eight Philadelphia schools by observing classrooms, meetings, and professional development sessions, and by interviewing teachers, principals, and other school officials. We interviewed district officials and civic leaders, and observed numerous meetings in which the reforms were debated, designed, and revised. We examined the SAT-9 test results and other indicators of system performance.

BEST COPY AVAILABLE

THE PLAN FOR TRANSFORMING PHILADELPHIA'S SCHOOLS

In 1995, with the support of the Annenberg Foundation, newly appointed Superintendent David Hornbeck launched Children Achieving, a ten-point reform agenda that promised what "no city with any significant number and diversity of students" had ever done before. Hornbeck boldly claimed that implementation of his plan would help "a large proportion of its young people achieve at high levels" (School District of Philadelphia 1995, p. i). The task was daunting. A special section of the *Philadelphia Inquirer* (1994) published just a few months earlier had painted a dismal portrait of the conditions in the school system. According to the *Inquirer*:

- Half the district's 220,000 students were from families on welfare.
- 136 of 238 schools were severely segregated.
- Over half of the city's public school students were failing to master basic skills. Fifty-one percent had failed the state reading test as compared to 13 percent statewide, and 50 percent failed the state math test as compared to 14 percent statewide. Seventy percent of African Americans and 75 percent of Latinos failed one or both parts of the state test.
- Forty-nine percent of ninth-graders failed to earn promotion to the tenth grade.
- On any given day one in four students was absent from class, and, in the average year, nearly one in four students was suspended from school.

To change these conditions and raise achievement, Hornbeck proposed an ambitious plan modeled after the reforms he had helped design for the state of Kentucky a few years earlier. The goals of Children Achieving were to transform the district into a school system characterized by

- high standards for all students
- accountability for results at all levels
- decentralization of authority from central office to schools and clusters of schools
- expanded professional development for teachers and administrators
- early childhood education for all children
- effective use of community services and supports
- adequate technology, instructional materials, and facilities
- strong public engagement
- adequate and effective use of resources
- comprehensiveness, or "Do all of the above at once"

The theory of action – the chain of logic about how these ten components would lead to improvements in teaching and learning and hence improved student performance – was not made explicit. Teams of central office staff, school staff, and community members developed plans that set forth the details for implementing the reform. Based on examination of these plans, other statements made by Superintendent Hornbeck and other district officials, and the actions taken by the district after the plan's adoption, the evaluators described the plan's theory of action as follows (CPRE et al. 1996):

If the district

- works with the schools and the community to set high academic standards for student achievement;
- aligns assessment with those standards;
- establishes an accountability system that offers strong incentives;
- delegates more authority over school resources, organization, policies, and programs to the schools;
- monitors equity throughout the organization; and
- builds public understanding and support for reform;

and if central office and the clusters

- provide guidance and high-quality support to schools and small learning communities;

then the teachers and administrators of the Philadelphia schools, in consultation with their

communities, will be motivated to develop, adopt, or adapt instructional technologies and patterns of behavior that will help all children reach the district's high standards.

The superintendent accepted this as an adequate summary of his reform ideas.

This theory of action highlights some key beliefs and values – articulated in district documents, in speeches made by Superintendent Hornbeck and other leaders, in interviews with district staff, and in discussion at policy meetings – that underlie the reform. They included:

All children can learn to high standards. The central tenet of Children Achieving was that "All children can learn, and 'all' means 'all.'" "All" included classified students, second-language learners, and all students at risk of poor performance.

The focus should be on results. To Hornbeck and his supporters, results were what mattered; how they were achieved was, at least in theory, less important.

Equity is paramount. The school district must be an advocate for the poor children it serves. Equity – of academic expectations, learning opportunities, and achievement outcomes – was a paramount objective.

School personnel need autonomy to meet the needs of their students. The theory of action and the work plans were based on an assumption that those working closest to students knew what was best for them, and wanted and needed the freedom and authority to act on this knowledge. Hence, central authorities in the district should not prescribe the means to achieve the goals lest they inhibit decisions and action by school staff.

Strong incentives are necessary. To spur action at the "cluster"[2] and school level, strong incentives had to be developed. Incentives included rewards and sanctions for performance as well as for adopting particular strategies or behaviors.

Do it all at once. Reform in all aspects of the system had to occur simultaneously and immediately to achieve significant results.

Not all of these beliefs and values were consistent, nor were they given equal weight or consistently apparent over the course of the reform. But the emphasis on being comprehensive and systemic (do it all at once) was strongly held throughout the reform and presented a challenge to formative evaluation. The district leadership's belief in comprehensive reform was so strong that, even though formal and informal opportunities for feedback to district policy-makers were frequent, comments about the confusion created in the schools by the simultaneous rollout of multiple reform activities were generally disregarded. District leaders felt that the benefits of integrated reforms would be lost if their implementation was sequenced over time.

We will return to these beliefs and values throughout this chapter, demonstrating how they shaped policy development and reform implementation; relationships among the schools, clusters, and the central office; and roles and decisions of central office leaders.

The Critical Drivers of the Reform

The critical drivers in the theory of action were standards, accountability, and decentralization. The reformers believed that these policy levers would energize the district and motivate staff and students to work towards higher performance.

Standards

Content standards were a cornerstone of Children Achieving.[3] Beginning in early 1996, teams of teachers were assembled to write standards in all subject areas. By late August 1996, draft standards for reading/English, language arts, mathematics, science, and the arts were distributed to teachers. Content standards in the social studies, health/physical education, and world languages followed soon thereafter.

Each set of content standards outlined the knowledge and skills that Philadelphia students were expected to acquire, with benchmarks, or performance standards, defined at the fourth, eighth and eleventh grades. In addition to requiring significant changes in curriculum, the standards also asked

2. Clusters were created under Children Achieving and are the district's intermediary organizational unit between the central office and the schools. There were twenty-two clusters, each organized around a comprehensive high school and the elementary and middle schools that feed into it.

3. Initially, performance and opportunity to learn standards were also envisioned, but they were never fully developed.

Philadelphia's accountability system was designed to assess schools' performance annually, and to reward progress or sanction decline every two years.

teachers to infuse "cross-cutting competencies" – skills and values in technology, multicultural competence, and communication – in all content areas.

It is important to note that Philadelphia's content standards did not specify a curriculum. Though they superseded the previous administration's "Standardized Curriculum," which prescribed a scope and sequence by grade level, the content standards simply defined broad parameters within which teachers and principals were expected to design their own curriculum. While consistent with the theory of action, this turned out to be an Achilles heel of systemic reform in Philadelphia.

Accountability

Philadelphia's accountability system, the Professional Responsibility Index (PRI), was designed to assess schools' performance annually, and to reward progress or sanction decline every two years. The PRI was made up of five indicators: student performance in reading, mathematics, and science as measured by the Stanford Achievement Test, ninth edition (SAT-9); a combined measure of teacher and student attendance; and the promotion rate (for elementary and middle schools) or the persistence rate (for secondary schools). These indicators were combined mathematically into the PRI, which provided each school with an annual score and with improvement targets.

The baseline year for the PRI was 1995–1996. Biennial targets were set for every school that assumed each school would make consistent, linear progress from its own baseline. New baselines were calculated every two years. The ultimate goal was for all schools to achieve or exceed a score of 95 on the PRI (out of a possible 120 points) by 2008. This score would mean that the average child in every school was achieving proficiency in the core subjects as measured by the district tests. Schools that met or exceeded their biennial targets were to be rewarded with cash; schools that did not meet their targets would receive assistance. If these interventions failed to bring improvement, schools faced reconstitution, the ultimate sanction in this scheme. Although two high schools were identified for reconstitution in 1997, this sanction was not employed in the first five years of Children Achieving.[4]

The accountability plan included the development of more challenging promotion standards for students at grades four and eight and new end-of-course examinations for core high school courses, but these were to be phased in beginning in 2000. New curriculum-related assessments had to be developed, and the superintendent and board of education, acting on the belief that students should not be subject to sanctions without appropriate supports, made the adoption of promotion standards contingent on securing additional funding for student supports such as an extended school day and summer school. The development of the new assessments began in 1999 and they were field-tested in the spring of 2000. However, when the deadline arrived for their adoption as promotion requirements, the district lacked the resources to provide the promised supports, and implementation was further delayed. This action kept intact the principle of linking pressure with support; however, the accountability provisions in Philadelphia ended up unbalanced for five years, falling heavily on teachers and school administrators but initially less so on the students whose effort was required to improve achievement.

4. Soon after the 1996 baseline scores on the SAT-9 were announced, the school district also announced its plans to reconstitute two high schools. The Philadelphia Federation of Teachers was outraged and charged that the reconstitution plans had been made without the appropriate consultation and before setting mutually agreed-upon criteria. An independent arbitrator agreed with the union and the reconstitution plans were abandoned, but not without cost. The episode seriously disrupted the two high schools marked for reconstitution (the principal of one of the schools had her car vandalized and was the subject of threats for her support of Hornbeck's plans) and embittered an already tense relationship with the teachers' union.

Decentralization

Along with standards and accountability, the other primary driver of the reform was decentralization. As conceived in Philadelphia, decentralization had four major components:

- *Small learning communities.* Small learning communities were intended to improve the conditions of teaching and learning in all schools, to strengthen relationships among teachers and between teachers and students, and to be the primary vehicle for improving instruction. They were subunits of schools, typically including four hundred or fewer students across several grade levels as well as the teachers responsible for their instruction. Some of Philadelphia's high schools and middle schools had voluntarily experimented with similar strategies prior to David Hornbeck's arrival, but small learning communities had not spread across the district until mandated as part of the Children Achieving reforms.
- *Local School Councils.* Each school was expected to establish a Local School Council (LSC) comprised of teachers, parents, the principal, and, at the secondary level, two students. The councils were to oversee school policies, review the budgets of small learning communities, and develop action plans to involve parents and communities in their schools to help improve student achievement.
- *Clusters.* The district was divided into twenty-two clusters. Cluster offices were to work directly with schools in support of reform. In Hornbeck's view, they were the "engines of change." Cluster offices had staffs who worked with a comprehensive neighborhood high school and the middle and elementary schools in its feeder pattern. The size of the cluster offices varied from a half dozen staff to over twenty depending on the cluster's capacity to raise external funding. The first six clusters were established during spring 1995 and the remaining sixteen in fall 1996. Clusters were expected to play a catalytic role in school improvement, guide and monitor the implementation of the reform agenda, provide focus for improvement initiatives, supervise principals, energize the schools, and mobilize resources. They also were expected to provide professional development, coordinate social services for schools, and strengthen K–12 articulation.
- *A streamlined central office.* The blueprint for the Children Achieving initiative clearly stated that the functions of the central office would be limited; it would "set standards, assess progress, monitor for equity, and act as a guide and provider of resources and support" (School District of Philadelphia 1995, p. iv.). This newly streamlined central office would give schools and clusters the freedom to make instructional decisions and put in place an infrastructure to help them make good decisions.

Through these changes in the structure and organization of the school district, the architects of Children Achieving hoped to increase the commitment of educators and parents and to improve productivity by reallocating power and resources and by reducing the isolation of teachers and school administrators.

Supports for Reform

District leaders recognized that Philadelphia's teachers and administrators would need new knowledge and new tools to implement these reforms and that they would need considerable support to do it well. They devised supports to help schools implement the standards and meet their performance targets. These included expanded professional development for teachers, curriculum frameworks, and family and community supports for students.

Professional Development

The Office of Leadership and Learning (OLL) was charged with developing and implementing a plan for professional development for administrators and teachers. It also was responsible for identifying and disseminating "best practices" – research-based reforms aligned with the new content standards.

The Teaching and Learning Network (TLN) was part of the OLL and served as the professional development arm of the district. However, TLN coordina-

tors and facilitators were based in the cluster offices and took their primary direction from cluster leaders, rather than from district staff. The number of TLN staff in a cluster varied from a few to over twenty as a result of differential district funding of the clusters and their varying capacity to obtain external funding. The TLN staff were expected to help schools and teachers implement district mandates and programs and to support the improvement initiatives of the schools and clusters. They offered workshops and classroom coaching for teachers. Largely recruited from the Title I program, most of the TLN staff had elementary backgrounds and lacked the content knowledge needed to help middle and high school teachers. Moreover, the pressures of implementing many changes simultaneously meant that they were usually forced to provide broad coverage rather than intense support. And the high levels of teacher turnover in some schools meant that their support often had to be focused on new teachers.

In response to concerns from teachers that the standards appeared divorced from content and pedagogy, making it difficult for them to develop curriculum, the district sponsored summer content institutes – weeklong professional development sessions in each core discipline linked closely to the standards – in the summer of 1997, a year after the standards had been adopted. They were well received by teachers, and participation increased dramatically over the course of the reform.

Districtwide Curriculum and Instructional Initiatives

Curriculum Frameworks. Developed in spring 1998 in response to teachers' requests for more guidance on how to implement the district's standards, the Curriculum Frameworks offered examples of instructional activities, units of study, and assessment tools for the standards in each subject area for every grade. They did not provide a coherent curriculum, however, nor did they provide a scope and sequence. While teachers found them useful, they continued to ask for more specific curricular guidance and for instructional materials related to the standards. Thus, by the fifth year of Children Achieving, central office staff were beginning to develop a districtwide curriculum.

Other Instructional Initiatives. Two other districtwide initiatives provided materials and sustained professional development to Philadelphia teachers during the Children Achieving era. One was the Philadelphia Urban Systemic Initiative, a five-year (1995–2000) systemic change effort funded by the National Science Foundation (NSF) with the goal of raising mathematics and science achievement for all students. Its strategies for change included providing teachers with intensive curriculum-based professional development and standards-based materials, promoting effective programs approved by NSF, and creating and supporting a network of teacher leaders. Interviews and observations of teachers indicated that this strategy was quite effective. Teachers who participated in USI activities were more supportive of the reform and much more likely to be implementing standards-based practices in their classrooms. This might have been because they were receiving curriculum units and content-specific professional development, neither of which was available to other Philadelphia teachers in the early years of the reform.

The second initiative, Early Balanced Literacy, was undertaken by the district in 1998 to ensure that children would leave the primary grades with a strong foundation in reading and writing. In the early years of Children Achieving, a number of elementary schools adopted or developed early literacy programs using a balanced phonics/whole language approach. Based on the success of these schools, the central administration made early literacy a districtwide focus and provided participating schools with materials and professional development as well as literacy interns to reduce class size. The Annenberg Foundation also provided additional funds to support early literacy programs in several clusters.

Family and Community Supports

The Family Resource Network was led by cluster staff and included school personnel such as nurses, guidance counselors, and teachers. It sought to strengthen student support services by mobilizing and coordinating community-based agencies and

direct service providers. Together with school personnel, they were expected to provide the "safety nets" that so many poor children need. The superintendent also proposed that city and private agencies work together to ensure that all students entered school ready to learn by expanding early childhood opportunities. Although the district successfully implemented full-day kindergarten systemwide, the envisioned early childhood initiative never got off the ground.

Research as a Support for Reform

The evaluation of Children Achieving was as comprehensive as the reform itself. Asked to play the roles of both formative and summative evaluators of the initiative, researchers from the Consortium for Policy Research in Education and Research for Action were expected to track the implementation of all the major components of the reform and to document their impact on classrooms and schools and student outcomes. Between 1995 and 2001, the research team interviewed hundreds of teachers, principals, parents, students, district officials, and civic leaders; sat in on countless meetings at which plans were designed, debated, and revised; observed the implementation of reforms in classrooms and schools; conducted and reported on two systemwide surveys of teachers; and carried out independent analyses of the district's test results and other indicators of system performance.

Regular oral feedback and periodically released written reports were provided to district leaders. Indeed, one of the supports for reform was our ongoing feedback on its progress in the schools. This was particularly true in the first several years of reform, when the role of the Children Achieving Challenge (the fiscal agent for the initiative), was particularly strong. For example, the initial report on the Family Resource Network led to a reorganization of the work of the Network; broader efforts to inform teachers of the new roles that school-based student-support professionals were supposed to play; and an emphasis on the collaboration necessary among student-support and instructional staff in schools. The district was also advised on needs that were not met by the original strategic action design, which led to the creation of summer content institutes and the development of curriculum frameworks. And problems were noted in the design that had been generating resistance or blocking implementation, such as the recruitment of TLN staff and the treatment of maternity leave as absenteeism in the accountability system.

As the Challenge's role weakened, our critical input had less impact. District leaders continued to seek input on specific issues that needed attention and were happy when the findings shed a positive light on an element or outcome of the reform. However, when the feedback challenged a strongly held belief or a core aspect of the reform, the information seemed less welcome. District leaders always listened to the findings or read them carefully, but, over time, fewer and fewer recommendations were implemented.

IMPROVEMENT AND CHANGE OVER THE COURSE OF THE REFORM

In this section we describe the impact of the reform program under Children Achieving on curriculum, instruction, and student performance. In the five years of Children Achieving, student test scores in Philadelphia as reported by the district rose significantly, although unevenly.[5] Gains were greatest in the first two years of the reform; they began to plateau in the third and fourth years. In the baseline administration of the SAT-9 tests in 1996, averaging across all subjects and grades, only 29.9 percent of the students tested scored at the basic level or above. The percentage scoring at this level rose to 41.9 on the 1999 tests. Table 1 presents the test results by subject and grade level. While gains were made in all subjects and at all levels, the improvement was most consistent in the elementary and K–8 schools.

The improvements displayed in Table 1 are especially noteworthy given that Philadelphia also aggres-

5. Philadelphia began using the SAT-9 in 1995–1996, the first full school year of Children Achieving. This analysis includes data from spring testing in each year from 1996 to 1999.

sively promoted the testing of all students. Compared with other urban school districts in this period, Philadelphia had one of the most inclusive testing policies, testing many special education students and English-language learners. From 1996 to 1999 the proportion of eligible students tested increased by 16 percent.[6] Since the students who were untested in the initial year of Children Achieving were likely to be lower achievers on average than those who were tested, the increased participation in the testing program undoubtedly acted as a drag on districtwide performance. Yet test scores rose significantly, in spite of the increased inclusion of these lower-performing students.[7]

However, there was one feature of this analysis that needed to be examined carefully. The major gains reported in student performance on the SAT-9 were based on an increase from year to year in the *percentage of relatively high-achieving* students (i.e., students who scored at basic or above) with respect to the *whole school*, including untested students. An analysis based only on the *average* performance of tested students (i.e., not including untested students) showed gains that were still statistically significant but not quite as dramatic. The reason for this apparent discrepancy is the influx of formerly untested students. As a higher and higher proportion of students were tested, both the percentage of students scoring at or above basic *and* the percentage of students scoring below basic increased in many schools, especially in middle and high schools, where the proportions of untested students were largest.

To illustrate, we provide this hypothetical example: In a school's first year of evaluation, 30 percent of the students were untested; 35 percent of students scored at or above basic; and 35 percent scored below basic. In the second year, after a campaign to include more students in the testing, only 10 percent of students remained untested. Most of the 20 percent of students that were newly tested in the second year scored below basic, but improvements in the rest of the test-takers moved a net of 10 percent of students from below basic to at or above basic. Therefore, the second year's figures would be: 10 percent untested, 45 percent at or above basic, and 45 percent below basic. While it was true that the *percentage of students performing at or above basic* increased from 35 to 45 percent, it was also true that the percentage of students performing *below* basic increased, so the *average* performance did not increase or increased more modestly.

Figure 1 compares these two ways of analyzing the Philadelphia SAT-9 data. Figure 1A shows the percentage of Philadelphia students scoring at or above basic on the SAT Reading test from 1995–1996 through 1998–1999, by school level. The percentages rose significantly from year to year. Figure 1B shows the average student performance as measured by the mean normal curve equivalent (NCE). Increases are still apparent, but they are smaller.[8]

The most robust gains were made in elementary schools, followed by middle schools. The average performance of eleventh-graders was flat over the course of the reform.

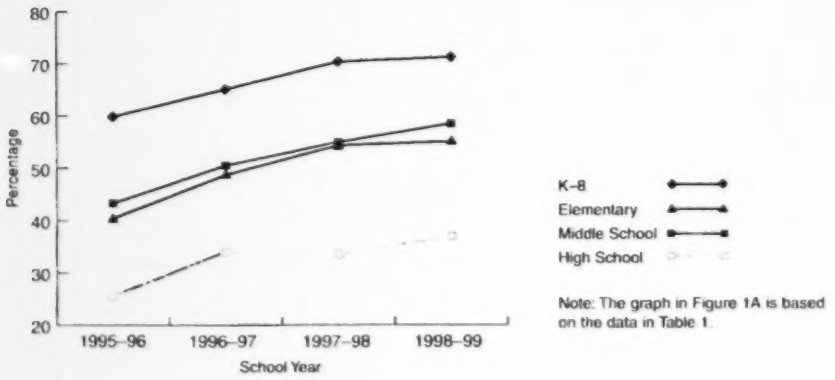
6. Under Children Achieving, schools that did not test all eligible students were penalized in the district's accountability system.

7. While these improvements in achievement were encouraging, it must be noted that the overall performance of students in the district remained low relative to other Pennsylvania districts. By 1999, the average eleventh-grade reading score was still more than 150 points below the state average, and the mean fifth-grade reading scores were nearly 200 points below the state average. (Data from Pennsylvania State System of Assessment: <*www.paprofiles.org*>.)

8. Raw scores on a given test are transformed into "norm-referenced" scores, of which the NCE is an example, to establish how an individual or group scored in comparison to a "norming group." For the SAT-9 that was used in Philadelphia, the norming group is a nationally representative sample of students at a particular grade level (e.g., U.S. eighth-graders). The NCE ranges between 1 and 99 and has a standard deviation set so that at three points – 1, 50, and 99 – the NCE would be equivalent to its corresponding percentile ranking, another type of norm-referenced score. In any test results reported using the NCE, including the SAT-9, we know that the average score is 50 and that half of the test-takers in the national norming group score above the mean and half below. NCEs are useful, like other norm-referenced scores, because they meet the statistical assumption of linearity and can be used in many linear analyses, unlike the percentages also reported in this chapter.

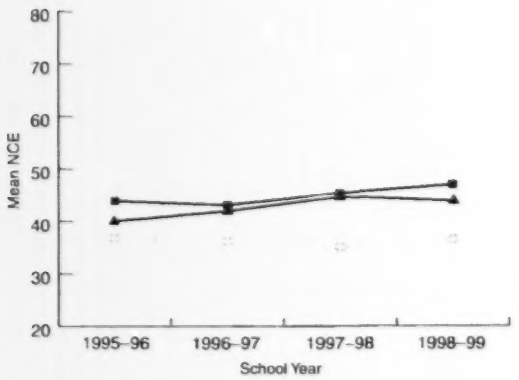
School Level	Reading					Math					Science				
	1995–1996	1996–1997	1997–1998	1998–1999	1999–2000	1995–1996	1996–1997	1997–1998	1998–1999	1999–2000	1995–1996	1996–1997	1997–1998	1998–1999	1999–2000
Elementary	40.4	48.7	54.4	55.1	57.7	32.2	41.9	45.3	47.8	49.1	37.1	46.4	52.5	52.2	61.5
K–8	59.9	65.1	70.4	71.3	n/a	40.4	48.5	53.7	54.5	n/a	42.1	51.9	57.6	59.3	n/a
Middle School	43.3	50.5	55.0	58.5	55.8	15.7	18.5	25.2	24.4	23.2	18.1	23.7	31.4	30.2	27.4
High School	25.6	34.0	33.6	37.0	40.7	11.5	13.8	15.9	15.2	16.4	4.9	8.0	8.4	12.0	19.1

Table 1. Percentage of Philadelphia students scoring at or above basic on the SAT-9, by subject area and school level, 1995–1996 through 1998–1999



Note: The graph in Figure 1A is based on the data in Table 1.

Figure 1A. Percentage of students scoring at or above basic on SAT-9 reading



	1995–1996	1996–1997	1997–1998	1998–1999
4th Grade	40.1	42.0	44.7	43.9
8th Grade	43.9	43.1	45.3	46.9
11th Grade	36.9	36.2	35.2	36.6

Note: normal curve equivalent scores were not available to us for the year 1999–2000. This figure shows data through the 1998–1999 school year.

Figure 1B. Average SAT-9 reading scores as measured by the mean NCE

Figure 1. Varying interpretations of improvement

BEST COPY AVAILABLE

Attendance, Promotion, and Persistence

Statistically significant gains were also made on all of the other indicators of school performance. The percentage of students in attendance for 90 percent or more of school days increased by three percentage points from 1996 to 1999; and the percentage of staff attending 95 percent or more of school days increased by over 6 percentage points. Persistence (on-time graduation) and promotion rates also increased significantly but less rapidly.

Relationship of Achievement Gains to Reforms

How can we account for these changes in performance? To what degree are they related to the implementation of the reforms? What aspects of the Children Achieving reforms seem to account for the achievement gains? Were the gains largely a response to the high-stakes accountability system? This section uses qualitative and quantitative evidence to address these questions.

As we have seen, Philadelphia elementary students made the most consistent gains on the SAT-9. Our quantitative[9] and qualitative research suggests that three factors contributed to these gains in elementary schools:

- test preparation
- focus on literacy programs in the primary grades
- development of strong professional communities in the schools

Test Preparation

Our qualitative data indicate that, in all likelihood, improvement in student achievement at all levels was to a significant degree the result of intensive test preparation and teachers' increasing familiarity with the content and format of the test.[10] While teachers initially saw inadequate curriculum guidance as a serious problem, the preeminent role of the SAT-9 test quickly shifted their attention from the content standards to the content of the test. Various forms of test prep were observed at all levels and were the most common instructional response to the reforms. Not all of this test prep was bad. Some of what might be considered "test preparation" were in fact educationally sound improvements in the curriculum to reflect the expectations in the test. These included an increase in writing assignments and higher standards for student writing. However, some of the test prep observed in Philadelphia was of the "drill and kill" variety. Teachers used materials such as Harcourt-Brace's KeyLinks workbooks to develop students' test-taking skills and familiarize students with the test.

Furthermore, the same form of the SAT-9 was used each year. While the district made an extraordinary effort to keep the test secure, researchers found copies of the test in the schools, and some teachers were familiar with the open-response questions. This problem arose again when the district field-tested its new fourth- and eighth-grade promotion tests and end-of-course exams. Copies of these new exams were readily available in the schools after they were administered, and the same tests were used in the subsequent year. We cannot say what effect the test security problems had on performance, but it is logical to assume that there would be some effect.

Focus on Literacy Programs in the Primary Grades

While at least some of the achievement gains can be attributed to the test preparation activities, the evidence suggests that instructional improvements also

9. Using hierarchical linear modeling (HLM), we examined the relationships between teacher and school characteristics, measures of professional community, measures of reform implementation, and growth in test scores over four years (1995–1996 through 1998–1999). The analysis showed that both the poverty level of students and the degree of professional community in a school were directly related to its growth in achievement in fourth grade. Interestingly, poverty did not depress growth in achievement. Schools with the highest concentration of poor students actually improved faster than schools with lower concentrations of poverty. But poverty was also a significant predictor of the baseline scores, so these poorest schools also had lower baselines and therefore more room for improvement. The only measure of professional community that was significantly related to growth in achievement was teacher collaboration: schools with greater teacher collaboration experienced higher rates of growth in achievement. The measures of reform implementation were not significant predictors of achievement growth in our HLM models, but subsequent analyses using logistic regression did reveal significant relationships between small learning community implementation and school conditions, as well as between small learning community implementation and professional community measures. For a more detailed explanation of this analysis, please contact the authors.

10. Efforts to create a quantitative measure of the level of test preparation in schools were not fruitful.

played a role in the gains and in their unevenness. In particular, qualitative data indicate that Philadelphia's focus on early literacy paid off in the primary grades. Classroom observations showed that teachers in the early grades increasingly used a balanced approach to teaching reading and writing, cooperative groups, and an emphasis on drafting and revising.

In contrast, Children Achieving did not offer middle and high school faculties equally specific or effective approaches to instruction. For the most part, middle and high school teachers did not focus on one or two robust and substantive strategies for improvement in student achievement. They faced large numbers of students who lacked basic skills and who required considerable support and remediation. Many students were alienated and hard to engage in academic work. In response to district mandates, teachers tinkered with the structural arrangements, creating small learning communities, interdisciplinary curricula, project-based learning, and service learning. In middle schools, they created new curriculum tied to their small learning community themes. After five years, we judged most of this thematic curriculum work to be still at an early stage of development. It seldom involved students in rich intellectual work, nor was it informed by multidisciplinary perspectives. High school faculties expanded opportunities for students to participate in internships and service learning, but were less successful at making classrooms more challenging learning environments or stimulating deep changes in instructional practice.

Strong Professional Community

While our analysis found no direct relationship between the degree of implementation of the Children Achieving reforms and growth in student achievement, we did find that the relationship may have been indirect. We found that well-implemented small learning communities were connected with teacher reports of higher levels of professional community,[11] and that there was a relationship between the strength of the professional communities, positive school conditions, and improved student achievement in elementary schools. That is, our analysis suggests that the implementation of small learning communities was associated with higher levels of professional community and that higher levels of professional community were linked to improved student achievement (controlling for significant factors such as poverty). Given the limitations of our data, no causal relationship can be inferred, but the findings do suggest possible directions for future work in the district and for further research.

The data also suggest that, in some schools, strong professional communities and positive school climates preceded Children Achieving and offered fertile ground for the creation of small learning communities and substantive pedagogical change. This was the case in at least two of the elementary schools and one middle school of the twenty-one schools[12] where we conducted intensive, multiyear qualitative fieldwork.

While small learning communities appeared to provide some benefits, they also generated new problems. In schools in which teachers were permitted to choose which small learning community they wanted to join, inequitable distributions of teaching talent sometimes resulted. The hardest-working, most able teachers often chose to work together, leaving some small learning communities staffed by those more resistant to the reform or less motivated. Students in most high schools were able to rank the small learning communities by their "quality," evidence that a form of de facto tracking had emerged. In addition, pressure for students to take most of their classes within their small learning communities meant that students' access to the full curriculum varied. This was most obvious in science, mathematics, and foreign languages, where not enough specialists were available to provide equivalent staffing in all small learning communities. A related problem was the variation in the quality of curriculum and curriculum implementation across small learning communities that followed the demise of middle and high school academic departments. Content teachers were in many cases on their own, with little access to support from peers in their fields.

11. We measured "teacher professional community" through our teacher survey. It encompasses items that describe the dynamics among teachers, teacher-principal relations, teacher collaboration, and principal leadership (see Tigue, Wang & Foley 2002).

12. This sample of twenty-one schools included eleven elementary schools, five middle schools, and five high schools.

LESSONS LEARNED FROM CHILDREN ACHIEVING

Children Achieving was a significant reform effort in both the amount of political capital expended and the investment made, but its effects were modest at best, and it was not sustained. What lessons can be drawn from this experience? Here we offer the insights we have gained from our five years of studying the reform in Philadelphia. They are not new insights, but they bear repeating because of their importance to the success of school reform in any school district, and to the role evaluators might play in the implementation of a reform strategy.

The Importance of Context

Philadelphia was not a wealthy city in the 1990s. Its population had decreased dramatically from the 1970s to the 1980s, and so had its middle class tax base. During that decade, the total population of the five largest U.S. cities – Philadelphia among them – decreased by 9 percent, while the population living in poverty grew by 22 percent (Wilson 1987, p. 46). When David Hornbeck began his tenure as superintendent, the city was still recovering from a serious economic crisis.

With that history, the city refused to provide significant additional resources for Children Achieving, arguing that it had "stretched its taxing ability to the limit" (School District of Philadelphia 1998, p. 26). But the full implementation of Children Achieving required significant additional funding, more than the $30 million generated annually by the grant,[13] and its design assumed that more funding would be forthcoming. In launching the initiative, Superintendent Hornbeck and his supporters took a calculated risk that the Annenberg Challenge grant would enable them to improve the performance of the system, and that evidence of improved performance would generate the political will needed to obtain increased funding through the city, the courts, or the legislature.

By 1997, the superintendent, the board of education, the city council, and the mayor all agreed that the Commonwealth of Pennsylvania was not upholding its fair share of the costs of educating Philadelphia's students. But state officials did not see it that way.[14] They believed that funds were being used inefficiently in Philadelphia and that the district's teacher contract was a major obstacle to improvement. In their view, better management and a better contract were prerequisites for additional state funds. The school district and the city used many strategies to induce the state to provide additional funding – multiple lawsuits, political brinksmanship, public scolding – but to no avail. The annual fiscal crisis of the school district became one of the few constants of the Children Achieving reform era.

Without new financing from the state, per pupil funding in Philadelphia remained well below what was spent in the surrounding areas. In 1997, Philadelphia spent $6,812 on each public school child. When compared to wealthy suburban school districts such as Jenkintown, Lower Merion, and Radnor, the gap was as much as $5,443 per student (School District of Philadelphia 1998, p. 11). Teacher salaries were also higher in suburban areas. Starting salaries in the suburbs were more than $3,500 higher than starting salaries in Philadelphia and maximum salaries were more than $9,000 higher (ibid., p. 29). Average teacher salaries in Philadelphia also fell below statewide teacher salary averages.

According to the school district, expenditures on administration declined during the Children Achieving era, although to make that claim they had to count the cost of the Teaching and Learning Network (TLN), professional development specialists based in the cluster offices, as an instructional expense. Critics who saw the TLN staff as an "administrative" cost contended that expenditures on administration actually grew over the course of Children Achieving.[15]

13. Though a significant source of discretionary funds, the $30 million from the Annenberg Challenge grant was equal to only about 2 percent of the $1.5 billion annual budget.

14. The funding that Pennsylvania provides to each school district currently is based on a funding formula which takes into account the number of pupils, the special needs of the district, its ability to raise local taxes, and other factors. However, the state froze the formula in 1993, which meant that state aid to the district did not rise in response to increases in enrollment and poverty in Philadelphia. In per pupil dollars adjusted for inflation, the real value of state education funds coming to Philadelphia annually between 1993 and 1998 actually decreased by 5.9 percent (see Century 1998).

15. See, for example Snyder 1998; Kirsch 1998.

There is no doubt that the number of staff assigned to the central office was smaller at the end of the reform than it was before it, but with over two hundred staff assigned to the cluster offices, school personnel felt that there was more bureaucracy, not less.

Inadequate resources limited the school district's ability to provide time for teachers and other personnel to receive professional development, to develop curriculum, and to work with colleagues. They also hampered the district's ability to hire the most qualified personnel. Teachers, in particular, had to make a real commitment to urban education (or be unable to obtain a job in the suburbs) to accept the lower starting salary in Philadelphia, a salary that was further reduced by the city tax on wages. Scarce resources also limited the ability of the school district to provide up-to-date curriculum materials and technology.

The Need to Build Constituencies and Partnerships

Social capital is a product of relationships among people. For example, a group of people who trust each other has a form of social capital. All other things being equal, a trusting group is more likely to succeed at a given task than a group whose members do not trust each other (Spillane & Thompson 1997). In Philadelphia, the limited social capital in the school district and inadequate efforts to build stronger constituencies for reform affected the implementation of Children Achieving. In particular, the culture of the district, its history of reliance on line authority, and the relationship of the central administration with its potential partners, including the Philadelphia Federation of Teachers, its own cluster leaders, state officials, and the business community, all affected the supply of social capital.

The Philadelphia Federation of Teachers. The school district's relationship with the Philadelphia Federation of Teachers (PFT) suffered over the course of Children Achieving and was characterized by mistrust on both sides. In a sense, the school district administration was at war with its own teachers. The PFT strongly objected to key components of the reform plan, particularly to its accountability provisions. They objected to spending money on cluster staff when schools were understaffed. The leaders of the PFT felt that Children Achieving was a threat to the union and to hard-won work rules in the teacher contract. Tensions were highest when the school district administration attempted to reconstitute two high schools; as a result of union objections, plans for the reconstitution were ultimately halted by an independent arbitrator who ruled that the district had failed to engage in the necessary consultation with the teachers' union. To the PFT leadership, the reconstitution attempt was just one example of the Hornbeck administration's pattern of excluding them from the decision-making process.

School district leaders, for their part, told us that the PFT representatives were invited to meetings about relevant policy areas, but that they either obstructed the meetings they attended or never showed up. Central office leaders felt that the PFT leadership was adversarial and unreasonably attached to the unproductive rules and regulations of an antiquated contract, and that the PFT had the interests of teachers, not children, at heart. In our estimation, both groups shared the blame. In four years of meeting with and interviewing central office staff and PFT representatives, we seldom heard positive comments from members of either group about the other and frequently encountered distorted interpretations of the other party's motives.

The acrimony on both sides of this relationship made progress difficult. The school district and the PFT were unable to agree on contractual changes that would have supported Children Achieving, especially in the area of decentralization. School communities were not permitted to select their own principals and staff, as Children Achieving advocated, and there was conflicting language about local school councils in the Children Achieving plan and the PFT contract.[16] Additionally, the failure of the school district to gain

16. The language of Children Achieving called for 35 percent of households to vote to determine council membership; in the teachers' contract the provision was for five parents to be selected by the Home and School Association. Additionally, Children Achieving called for two-year terms for parents, while the contract outlined one-year terms for teachers. See Christman 1998.

concessions from the PFT undermined its credibility with a number of stakeholders, particularly principals and the business community.

The tension between the PFT and district leaders also made the job of evaluation more difficult. PFT leaders viewed our frequent meetings with district officials with suspicion. Though we were providing formative feedback to the district, often mentioning the need to develop better relations with the Teachers' Union, the PFT questioned our impartiality. In turn, our willingness to hear out the PFT's questions and objections to the reform, and to seek their cooperation in administering surveys of teachers, made some district leaders apprehensive, and may have reduced our access to them in later years of the reform.

Cluster Leaders. In addition to an antagonistic relationship with the leadership of the Philadelphia Federation of Teachers, the school district often alienated its own cluster leaders, who were crucial to reform implementation. Cluster leaders were, for lack of a better term, regional superintendents, who were supposed to improve and align instruction across a feeder pattern of schools and lead and support local professional development and community engagement. They were members of the superintendent's cabinet, which also included key central office leaders.

With the addition of twenty-two cluster leaders, the cabinet ended up being a group of about fifty people, a size that was ill-suited for collaborative work. Cluster leaders came to describe cabinet meetings as the place they came to talk about decisions that were already made by central office staff. Cabinet meetings were also one of the few forums they had to air their grievances and, as a consequence, central office staff often felt "ganged up on" by cluster leaders.[17] In one particularly contentious meeting, for example, cluster leaders were upset that more information was not available as to how they would finance and organize summer school programs, scheduled to begin about three months from the time of the meeting. They made little effort to hide their anger and hostility.[18]

This tension arose in part because of conflicting ideas about the cluster role. Whereas some central office staff saw clusters primarily as vehicles for informing the field about new aspects of the reform, cluster leaders felt they should have the autonomy to determine the means of improving performance. In interviews, they frequently complained about central office mandates and their lack of influence over policy. In the following excerpt from our field notes, a cluster leader illustrates this point:

> Part of the challenge I have had as a cluster leader is to keep the central office away from me, so I can allow my people to develop their responsibilities. Downtown keeps adding more [stuff] to our plate . . . . Let me give you an example. The central office wanted to change the special ed formula, which apparently they had been working on for months, but it wasn't shared with anyone [in the field]. When it was finally announced, parents went to the board and begged them not to let it happen. So the board then asked the school district what facts they have to support the change, so now we [cluster staff] have to do a lengthy survey. We have to identify one special ed kid per special ed classroom and review their [education plan], observe their classroom, interview the parent and teacher, and we have to do it all in four weeks. That's 75 kids for me because we have 75 special ed classrooms. The central office knew that they would ask us to do this in the summer, but they didn't actually ask us until a couple of weeks ago. This says to me that I have to put a hold on everything else I'm doing and do this. It takes away from your focus.[19]

The central office staff became aware of time and turf concerns and made efforts to seek cluster leader input and plan with respect to cluster schedules. Nevertheless, the unproductive relationship endured, and cost the district key support. In the 1999–2000 school year, the superintendent asked the cluster leaders to back him in a fight for funding from the state; they refused.

17. Personal communication, December 2000.

18. Field notes, March 24, 1999.

19. Field notes, April 1996.

State Officials and Business Leaders. The school district's relationship with state education officials, the governor, and the state legislature also was strained over the course of the reform. When Hornbeck became superintendent in 1994, there was a Democratic governor and Democratic majorities in both houses of the state legislature. He came to his position with strong backing from both Philadelphia's mayor and its business community. However, just three months into his administration, the political landscape in Pennsylvania and Philadelphia changed dramatically: the state elected a Republican governor, Tom Ridge, and Republican majorities in the state legislature who were committed to reducing government spending. Relationships between the state officials and the district were tested by the new governor's advocacy of vouchers, his refusal to grant the school district significant additional funds, and the superintendent's inflammatory rhetoric, alleging racism on the part of state officials in speeches and via a federal civil rights lawsuit against the state. When we interviewed state education department officials in the fall of 1999, their anger toward David Hornbeck was evident.

This antagonistic relationship between the state and the school district had effects on local constituencies as well. The strong backing of the business community for Children Achieving deteriorated as Hornbeck's battles with the state became more public. In addition, some of the superintendent's strongest supporters left Philadelphia as the major corporations headquartered in Philadelphia moved out. In civic organizations like Greater Philadelphia First, leadership shifted from executives of large national corporations to leaders of smaller, more local firms. The clearest sign of the fracture in the alliance between the business community and the school district was when Greater Philadelphia First – a coalition of Philadelphia business executives that served as the fiscal agent for the Annenberg Challenge – supported Governor Ridge's plan for school vouchers.

External Reform Support Organizations. From the beginning of Children Achieving, the school district had two primary external partners: The Children Achieving Challenge (CAC) and the Philadelphia Education Fund (PEF). Leaders from both groups were members of the executive committee and the superintendent's cabinet and played significant roles in the reform effort, especially in establishing and leading the seven work teams charged with developing goals, setting priorities, and creating annual work plans to implement Children Achieving.

The work teams served an important function by providing learning opportunities for the central office and cluster staff, one of the few such systematic opportunities they had. (We were part of a work team on the evaluation, which gave us many opportunities to discuss preliminary findings and get feedback on our research design.) At the outset of Children Achieving, systemic reform was a fairly new concept nationally and few anticipated the demands it would place on teachers and schools, let alone on the central office staff. In autumn 1996, for example, some central office leaders were still questioning whether standards were curriculum.[20] Most members of the superintendent's inner circle, particularly the PEF and CAC leaders, were knowledgeable and thoughtful about standards-based, systemic reform, but it took some staff at the central office a considerable amount of time to develop a deep understanding of the initiative they were helping to lead. The work teams were key to their understanding and led to important decisions about the role of the central office, the supports it would provide for the reform, and key contributions from the two external groups.

Unfortunately, the basis of the school district's partnerships with CAC and PEF were more personal than institutional. When the leaders of these organizations left Philadelphia to pursue other career opportunities, the organizations' relationships with the school district weakened. Collaboration continued, but the relationships were more marginal and less catalytic. The two leaders' departure further isolated the school district from key constituencies.

20. Field notes, November 12, 1996.

Importance of Focus

Another factor that made reform implementation difficult was the pressure on school staffs generated by the core belief that the whole system must be reformed simultaneously and immediately. The superintendent adamantly opposed piecemeal, incremental reform. He felt that the ten components of his reform plan were mutually supportive and had to be moved forward simultaneously. But "doing it all at once" created reform overload throughout the school district, from schools to the central office. School staffs were unable to focus their efforts around clearly defined and manageable instructional priorities. Cluster staffs were overwhelmed; they worked hard to win teachers' support and to assist them, but they were hampered by the sheer number of district initiatives and directives. Many clusters were unable to fully develop or implement their own reform strategies because so much time was spent promoting and disseminating information about new central office policies and programs that the schools were required to implement.

The volume of reform initiatives also overwhelmed many central office staffers, particularly those who were not among the superintendent's close advisers. The concern about overload was evident very early in the Children Achieving initiative. In an early policy meeting, when several of the superintendent's inner circle had left the room, one central office leader said, "We need to talk about priorities and make some tough, hurtful choices and let the chips fall where they may. We can't pretend any more that we can do it all." Another central office leader agreed, using the analogy: "We can't plow all the streets. Which ones are most important?" He suggested that focus should be placed on a group of schools or a few clusters.[21] But when the inner circle members returned and the other participants briefed them on what had happened in their absence, there was no mention of the concern about reform overload.

This reluctance to tell the superintendent and his closest staff about the difficulty of "doing it all at once" continued throughout the reform effort. In an interview two years after the exchange quoted above, a district leader told us:

> I've got to tell you something else. We are on innovation overload! As hard as it is for a superintendent in a large district, someone has to have the guts to say it....Everyone is tired....[Central office personnel] are having to learn something new all the time, we're rolling out so many competing forces. [Begins counting on his fingers.] We have the CSP. We have SLC. We have School-to-Career. We have service learning. We have multidisciplinary projects. And there is more to come. That is just one hand! We have judgments against us in federal courts that push us to make things not fall through the cracks....There's always a new priority.[22]

The urgency of "doing it all at once" created pressure on central office staff simply to "roll out" the reforms and move on to the next priority. There was little time to support or guide the reforms or to receive our feedback and review and revise policy. It is not surprising that, to schools and clusters, central office policy felt like unsupported mandates. The core value of "doing it all at once" increased the top-down mandates by the central office, conflicting with the core value of school autonomy.

As evaluators, we frequently tried to point out problems of this kind, but we were in an awkward position. There were many constraints on what we were able to do as witnesses to this lack of focus and still maintain the confidentiality of our informants and access to the system. Additionally, the lag time between the discovery of patterns of response such as the field's perceptions of the lack of focus – rather than merely individual instances – and their actual occurrences often made these findings less resonant with district leaders. By the time we shared our findings with them, they were often enveloped in the implementation of different aspects of the reform, a fiscal crisis, or another political battle. Moreover, the tension between the district and the teachers' union also made some of our information suspect to some in the administration. Because we had collaborated with the union on districtwide surveys, some believed that our repeated warnings about the effects of reform

21. Field notes, November 1996.

22. Field notes, December 1998.

burden, problems in sequencing, and the inadequacy of the support structures made us "apologists" for resistant teachers.

Importance of Reform Sequencing

One of the primary flaws in the implementation of Children Achieving was the sequence in which the district rolled out the reforms and supports. In order to capitalize on the momentum built up from the hiring of the new superintendent and the acquisition of the Annenberg funds and to fulfill the underlying belief in the need to "do it all at once," there was a rush to implementation. The district led with the pieces that were easiest to put in place. And, for strategic reasons, the district wanted to be able to demonstrate relatively soon that it was making gains in student achievement, to persuade the state, or the courts, that Philadelphia should receive increased funding so the reforms could be sustained. As a result, the new cluster organization, the new tests, and the accountability system were the first components to be implemented. The standards came in the next school year. The tools and supports needed by teachers to use the standards and prepare students for the tests came even later.

But to many, instituting accountability policies and other structural reforms before developing the infrastructure needed to support the changes in practice and services required to raise achievement was putting the cart in front of the horse. School personnel complained that they were being held accountable for performance targets before they had received the new standards, before all twenty-two clusters were in place, and long before the development of the Curriculum Frameworks offered a modicum of guidance and summer institutes offered teachers rich opportunities to examine their practice. All of these sequencing problems contributed to perceptions by teachers and principals that they were being asked to carry disproportionate amounts of the burden for improvement. They felt victimized by the ways in which the reforms were presented and rolled out. Faced with dire consequences, many teachers turned to sure and safe methods of instruction – drill and practice.

The urgency of "doing it all at once" created pressure on central office staff simply to "roll out" the reforms and move on to the next priority. There was little time to support or guide the reforms or to receive our feedback and review and revise policy.

Our evidence suggests that the teachers' concerns may have been valid. When the district finally put in place the instructional supports teachers had clamored for, those who were able to take advantage of them benefited from them and many improved their practice. A different sequence might have produced different results for the district.

Policy-Makers as Learners

The capacity of the central administration to support the Children Achieving reforms was an issue not only of financial resources, but also of human capital. With the exception of a few key leaders, knowledge about the substance of the reforms and how they fit together, and the expertise to implement them was limited, even in the central office. While many central office staff members were passionately committed to Children Achieving, some had only a superficial understanding of the reforms they were supposed to help schools implement and of the demands they made on teachers and school administrators. Much of what they were trying to implement existed only in theory prior to Children Achieving.

Poor personnel decisions and turnover in staff also limited central office capacity. The associate superintendent in charge of the initial development of the reform and the leader of the superintendent's transition team resigned in protest over the superintendent's insistence on promoting teacher accountability. A well-regarded central office leader was demoted for refusing to submit a resignation letter early in the

Embattled leaders attempting to implement an ambitious plan with key elements underdeveloped are not in a good position to make midcourse corrections on the basis of research information.

reform. A deputy superintendent retired. Over the course of the reform, there were three different leaders of the Office of Leadership and Learning, four directors of information technology (including two acting directors), three directors of the Office of Best Practices, three managing directors, and two directors of the Office of Curriculum Support.

Staffing turnover and ineffective leadership plagued the departments most directly responsible for providing support to the field – the Office of Leadership and Learning, the Office of Curriculum Support, and the Office of Best Practices. Staff hired to fill these vacancies were not, in general, compatible with other members of the leadership group and some gained reputations as "stallers" – people who put up obstacles to reform.[23] One central office leader admitted, "Central office personnel decisions have not been good ones."[24]

Researchers as Learners

Despite the district's stated commitment to evaluation, the experience in Philadelphia suggests that there are limits to the influence researchers can wield. Reformers working in highly politicized environments, in which the stakes are high and opponents are ready to take advantage of each mistake or sign of weakness, may be reluctant to admit to flaws in design or errors in strategy. And even if they are willing to make changes, they may be unable to do so. Embattled leaders attempting to implement an ambitious plan with key elements underdeveloped are not in a good position to make midcourse corrections on the basis of research information. They are committed; it is more expedient to blame the critics or the teachers who are struggling to implement their ideas than to admit error or come up with new plans on the fly.

To be sure, Philadelphia leaders did at times show a willingness to listen to criticism and a flexibility to adjust some of their plans. The district implemented well-regarded content institutes and then curriculum frameworks after teachers complained about a lack of curricular guidance. And the district delayed implementing its student promotion plan when funds for supporting services did not materialize. But for the most part, district leaders seemed unwilling to hear challenges to the fundamental tenets of Children Achieving, let alone act to change them. The response seemed to be to circle the wagons rather than build a new camp.

The lessons for evaluators seem clear. First, the same team of researchers cannot do both internal formative and public summative evaluation work in atmospheres as charged with conflict as Philadelphia was during the late 1990s. Providing formative feedback requires a high level of trust and a close working relationship with the leaders and designers of the reform. This cannot be established and maintained if the research team is also reporting to funders and sharing findings with potential critics of the reform. The high level of conflict generates a bunker mentality in which the norms are believing in the reform in toto and questioning core beliefs is tantamount to joining the opposition. Simply put, the people to whom we were charged with providing feedback were not in a position to accept constructive criticism from evaluators who also were asked to make summative judgments about the success of the reform.

Our limited impact was exacerbated by a timing problem. There was tension between the funders' desire for a broad-based view of all aspects of the reform and the researchers' interest in looking at the reform in depth. Therefore, the results from the research tended to arrive too late to be helpful to the reformers. They were either already deeply committed to a course of action or the circumstances had

23. Field notes, July 3, 1997.

24. Field notes, November 1999.

changed by the time the research team reported. In retrospect, the researchers should have conducted a series of shorter-term studies focused on points of potential conflict or tension within the system that could have provided rapid and useful feedback to the designers. This would have made it harder to study the "big picture" over time, but might have been a more useful contribution to the successful implementation of the reform.

CONCLUSIONS: LOOKING BACK

By the spring of 2002, Philadelphia public education was under a new regime. The state had exercised its authority to take over the financially troubled district, and a new five-member commission appointed by the governor and mayor had assumed control. They were moving forward with plans to contract with private firms and organizations to take over some of the city's low-performing schools. The fate of the other schools was not clear. Some of the central components of Children Achieving, such as the cluster system, the PRI, and the TLN, were gone. The fate of others – the standards, the curriculum frameworks, the small learning communities, and the Family Resource Network – was not clear.

City leaders seemed to be abandoning Children Achieving and its means for bringing about improvement. Advocates of the reforms argued that considerable progress had been made and claimed that inequities in state aid, resistance from the teachers' union, and declining support from the business community had undermined the possibility of even more progress. Critics of Children Achieving in turn pointed to fiscal deficits, increased expenditures on administration, the flattening of test scores, and the emphasis on test preparation as evidence that the reforms were seriously flawed. Because so many of the critics seem to be willing to dismiss the issue of inadequate funding, it is tempting to say that the truth lay somewhere in between. In our view, though, the critics seemed to have the more convincing arguments.

Without a doubt Children Achieving had offered a compelling set of ideas for school reform and had changed the nature of the debate over public education in Philadelphia. Central ideas such as the beliefs that results matter; that all children can learn at high levels, and that "all" means "all"; that everyone must be held accountable; and that professional development is a necessity generated a new set of expectations for local policy-makers. However, the leadership of the school district of Philadelphia paid too little attention to implementation lessons from the past when they crafted the reforms. They too often criticized teachers rather than attempting to win their support. They adhered to the dictum of the Children Achieving plan that everything had to be done simultaneously, which placed enormous burdens on teachers and principals. They assumed that teachers would embrace the reforms in exchange for more freedom to develop curriculum and more influence over school decisions. They were careless in the manner in which the reforms were sequenced. They put pressure on teachers before they provided supports, and they underestimated the difficulty of developing standards-based curriculum and instruction.

Philadelphia's policy-makers also lacked a clear vision of what was required to implement standards-based instruction at the central office, cluster, and school (administrative, teacher, and parent) levels. Capacity was lacking at all levels of the system, yet efforts to build it were sporadic and weak. Left without the necessary supports and feeling overwhelmed and overburdened, many teachers, principals, and administrators left the district seeking higher salaries and better working conditions outside the city, making implementation of the reforms even more difficult.

Individual schools also varied in their capacity for change, their professional cultures, and their reform histories. The experience in Philadelphia suggests that different reform strategies are needed for elementary, middle, and high schools. Each level of school brought different organizational issues, professional norms, and cultures to be addressed. Their past experiences with reform varied, as did the challenges they faced in motivating students and staff. However,

Children Achieving only offered a "one size fits all" reform strategy that was difficult to adapt to varying school contexts.

Ideally, researchers could have brought these flaws to light and worked with the district to adapt its strategies. Problems of design and implementation are precisely those that dispassionate observers, armed with knowledge about the experiences of other districts, could best help reformers address. But the researchers' influence was limited, and the midcourse changes the district made were not sufficient to fulfill its compelling rhetoric and promises of a new day in urban education.

Yet it would be inaccurate to say that nothing was accomplished in the five years of Children Achieving. First of all, there were some real gains, such as an improvement in elementary student performance and the introduction of full-day kindergarten and early literacy programs, although the gains were modest compared to the ambition and scope of the reform effort. Secondly, the reform effort raised expectations for Philadelphia's children; it forced citizens to recognize the sobering realities of public education in Philadelphia and to debate its future.

Finally, the Philadelphia reform experience presents us, as researchers, with the opportunity to look back at the difficult lessons about what brought Children Achieving to an end and to gain valuable insights that can help future reform efforts succeed. We learned much about the challenge of sustaining systemic reform and about the need for resources and support to match the complexity and ambition of the reform design. We also learned much about the issues surrounding the role of external evaluators in a reform effort. To the extent that educational stakeholders – students, teachers, school and district administrators, the community, and researchers – can help each other learn from constructive feedback, negative as well as positive, future reform efforts will be able to overcome the challenges of the past and help public schools fulfill their high hopes and ambitious plans to educate all students to high standards.

References

Century, J. A. 1998. *Citizen's Guide to the Philadelphia School Budget.* Philadelphia: Greater Philadelphia First.

Christman, Jolley Bruce. 1998. *Guidance for School Improvement in a Decentralizing System.* Philadelphia: Consortium for Policy Research in Education.

Consortium for Policy Research in Education, Research for Action, and OMG Center for Collaborative Learning. 1996. *A First-Year Evaluation Report of Children Achieving: Philadelphia's Education Reform.* Philadelphia: Greater Philadelphia First.

Kirsch, T. 1998. "President's Message: We Are Fighting Back – Together – and We've Got to Win!" *PFT Reporter.* Available at <*http://www.pft.org/pres798.html*>.

Pennsylvania State System of Assessment. Available at <*www.paprofiles.org*>.

Philadelphia Inquirer. 1994. "A District in Distress." Special Section, G1–G8 (October 23).

School District of Philadelphia. 1995. *Children Achieving Strategic Action Design.* Philadelphia: School District of Philadelphia.

School District of Philadelphia. 1998. *Realities Converge, Revisited: School District Sees Gains on Test Scores and Management Efficiencies, But Fiscal Crisis Is at Hand.* Philadelphia: School District of Philadelphia.

Snyder, S. 1998. "No Cash Crisis in Schools: Philadelphia's Problem Is Management, Says an Accounting Firm Hired by the State Legislature," *Philadelphia Inquirer* (November 13).

Spillane, James P., and Charles L. Thompson. 1997. "Reconstructing Conceptions of Local Capacity: The Local Education Agency's Capacity for Ambitious Instructional Reform," *Educational Evaluation and Policy Analysis* 19:2, 185–203.

Tigue, Erin, Aubrey Wang, and Ellen Foley. 2002. *An Analysis of Children Achieving on Student Achievement in Philadelphia Elementary Schools.* Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.

Wilson, W. J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* Chicago: University of Chicago Press.

CHAPTER 7:
BEYOND STANDARDIZED-TEST SCORES:
USING CASE STUDIES TO EVALUATE A REFORM STRATEGY

Ann G. Bessell, Jeanne Shay Schumm, and Okhee Lee
University of Miami

Elaine Liftin and Sisty Walsh
South Florida Annenberg Challenge

The South Florida Annenberg Challenge

THE SCHOOL DISTRICTS of Miami–Dade, Broward, and Palm Beach counties make up the South Florida Annenberg Challenge, which received a $33.4-million, two-for-one matching grant from the Annenberg Foundation in late 1996. The goal of the South Florida Annenberg Challenge (SFAC) was to increase students' academic achievement through principal leadership, teacher empowerment, and parent involvement in the tricounty region.

Between January 1997 and December 2002, the SFAC provided matching funding to "partnerships for public education," groups of three or more schools, a local business, and a parent or community organization. Each partnership was encouraged to develop innovative approaches to education issues, such as school readiness, technology, teacher training, and parent involvement.

The authors acknowledge the support of the South Florida Annenberg Challenge Executive Board and the Evaluation Committee, with special thanks to its chair, Steven J. Saiontz. They give particular tribute to the late Leonard Miller, SFAC Chair from 1997 to 2002, for his dedication to the education of children in the SFAC projects and in the South Florida schools.

SFAC made grants to eighty-nine partnerships that involved 378 schools, over 265,000 students, and some 5,000 teachers in the three districts. While most of the partnerships were in a single school district, several were regional partnerships across all three counties. All the partnerships focused on children who were potentially at risk for low academic achievement or school failure and who lived in some of the poorest neighborhoods of Miami–Dade, Broward, and Palm Beach counties. Many of these students were eligible to receive free or reduced lunch, were students of color, and/or had limited English proficiency. Thus, the SFAC projects were intended to benefit students most in need of educational opportunities.

The partnerships encompassed a wide range of goals, activities, and outcomes. Some focused on specific instructional strategies; others focused on curricular reform, on equal access to technology, on parent involvement, or on systemic change of education system. While the partnerships represented diverse interests, they pursued the common SFAC goals of improving student achievement and fostering systemic impact by promoting empowered principals, quality teachers, and parent involvement. Some projects were being newly implemented; others were in various stages of being expanded or replicated.

The SFAC program evaluation was designed to examine the effectiveness of the partnerships and to gain insights about community-based reform efforts. The evaluation focused on four aspects:

- examining the effectiveness of all partnerships,
- conducting in-depth case studies of selected partnerships,
- drawing lessons from effective partnerships, and
- providing recommendations to the SFAC board.

This chapter describes how case study research was used as the basis for the SFAC program evaluation. After providing a rationale for using case studies in the evaluation of school reform efforts, it continues with an overview of the evaluation design and framework guiding the evaluation. It then describes the case study methodology used from 1999 to 2001 and provides highlights of case study findings from the fourteen partnerships reviewed during that time. Finally, the chapter concludes with key lessons learned from the case studies and lays out future directions for the SFAC program evaluation.

RATIONALE FOR USING CASE STUDIES TO EXAMINE CIVIC-MINDED APPROACHES TO SCHOOL REFORM

Under increasing national attention, education reform has taken two distinctly different approaches: policy-minded and civic-minded (McDonald 1999). The policy-minded approach involves education policies and regulations mandated at the federal, state, and district levels. Currently, two strategies dominate the policy-minded approach: the *accountability* strategy, based on content standards and assessments in subject areas; and the *market commodity* strategy, designed to create better schools through supply and demand, with schools competing for students (e.g., charter schools, school vouchers).

An alternative to policy-minded reform is the civic-minded approach, based on locally developed public-private partnerships (Fitz & Gorard 2000, Bracey 1999). Like the policy-minded approach, the civic-minded approach may also employ accountability-based or market-based strategies. However, while the policy-minded approach depends on control from government and policy-makers, the civic-minded approach is based on close and inviting collaboration between public and private sectors in the community via an "inquiry-minded" process (Rallis & MacMullen 2000).

Although the civic-minded approach provides an alternative to policy-based school reform, its process and impact have not been adequately examined. In addition, research conclusions that rest too heavily on statistically analyzed standardized-test results may miss much of the formative details of the "process" of reform (Hoyle & Slater 2001).

Evaluation through Case Studies: A Good Fit for SFAC

Case study methodology was particularly suited to evaluation of SFAC (a prime example of the civic-minded approach to school reform) for three reasons:

To address the wide variety of approaches to school reform in SFAC. The range and diversity of SFAC initiatives made the design and implementation of survey and program-monitoring instruments particularly challenging. Case studies enabled the program evaluators to develop insights about the varied approaches to school reform encompassed in SFAC and to design instruments that would elicit meaningful data from key stakeholders. The instruments and results from case studies would then be applied to develop systemic evaluation procedures with all partnerships using surveys and monitoring reports.

To go beyond standardized-test results. At the outset of the evaluation effort, case studies were included as part of a multimethod evaluation to shed light on critical issues that arise when intermediary organizations, as agents of change, foster creative public-private partnerships and civic mobilization. The SFAC used statewide assessments of student performance in reading, writing, and mathematics as the key out-

come variable because there were no other consistent achievement measures across the range and diversity of SFAC partnerships. These high-stakes assessments are used to hold schools accountable, and the community at large is well aware of the impact of the assessments.

The case studies of partnerships were designed to look beyond standardized-test results as the sole measure of school reform by highlighting perceptions of key stakeholders about project implementation and outcomes. In addition, perceptions of key stakeholders from the case study partnerships provided insights to explain student-achievement outcomes.

The case studies of partnerships were designed to look beyond standardized-test results as the sole measure of school reform by highlighting perceptions of key stakeholders about project implementation and outcomes.

To provide timely feedback about reforming partnerships' efforts. School reform takes time, and student-level data reflecting the impact of partnership initiatives on academic achievement would not be available until the completion of the three-year life of an SFAC partnership. Fullan (2000) found that it takes three to five years of reform in elementary schools to show improvement in student performance on state assessments and up to six years in high schools. In the absence of meaningful student-achievement data in the partnerships' formative years, case study data enabled researchers to examine the degree to which the goals of each case study partnership were accomplished. The results, then, would be used to examine the outcomes of all partnerships and the impact of the SFAC project overall.

The Role of Case Studies in the SFAC Evaluation

Using case studies as the cornerstone of the SFAC evaluation served several major purposes for the evaluation. The case study procedures were used to develop the instruments for surveys and monitoring reports. The mostly qualitative results of case studies were used to help categorize and process the massive amounts of data from multiple sources and establish consistencies so that meaningful data from all partnerships could be found or generated and analyzed statistically. The case study results were also used to form a pathway for tracking early indicators of success until achievement data were available, as well as to provide explanatory mechanisms to evaluate the level of success of partnerships.

EVALUATION DESIGN AND FRAMEWORK FOR EVALUATION

The ultimate goal of the SFAC evaluation was threefold: to determine the degree to which each partnership (and schools within the partnership) were successful in achieving positive student-achievement outcomes, to determine the return on investment through cost-effectiveness analysis, and to draw lessons about effective partnerships and offer recommendations for sustainability and replicability of a civic-minded approach to education reform.

Evaluation Design

To evaluate the SFAC overall and selected partnerships in depth, a mixed-method research design was developed using parallel or simultaneous methodologies (Tashakkori & Teddlie 1998). This design was used because it provides in-depth evaluations of selected partnerships, addressing concerns about validity and reliability and, thus, the transferability of the findings to other partnerships. The use of parallel quantitative and qualitative methodologies also triangulated the data and strengthened the evaluation findings (Breen et al. 1998).

The qualitative methodology provided the basis for the case study investigations (Yin 1994) and

formed the framework that was used to examine the effectiveness of individual partnerships, the SFAC in general, and the impact of community-based reform efforts. Data were gathered from key stakeholders in all aspects of project participation, adding depth and breadth to the analysis. These findings provided evidence, validation, and the opportunity to learn about unexpected findings (Seale 1999). The qualitative methods, including individual and focus group interviews and observations of case study partnerships and monitoring reports of all partnerships, served as critical data sources for obtaining participants' perceptions. Documenting and analyzing these perceptions in a systematic way provided a better understanding of the "how" and "why" behind project implementation and outcomes (Morgan & Krueger 1993; Posner 1995; Rubin & Rubin 1995).

The quantitative methodology provided the primary data for assessing outcomes related to student achievement and for measuring the perceptions and satisfactions of key stakeholders (using surveys with all partnerships). This methodology used descriptive research to provide insights into the current status of each project and schools within each project and to gather baseline data for future research (Gay & Airasian 2000). In addition, this design was selected because it would allow for analysis of student-achievement outcomes and cost-effectiveness in relation to comparable schools, each school district, and the state.

Framework for Evaluation

The framework guiding the case study investigation was developed from an extensive review of the literature pertaining to school reform, business/school partnerships, public policy, and success indicators (which were also gleaned from cross-case analysis of earlier case studies). Information from these sources

A Partnership of Program and Evaluation

Ambassador Annenberg framed his Challenge with a belief in the potential of the public education system. Rigorous evaluation would be a necessary component of each "living" initiative as it evolved. Local evaluations of Challenge efforts needed to be purposeful as well as reflective, to balance accountability with flexibility in stimulating innovative and responsive local programs – no easy task.

Right from the start, business and philanthropic supporters of the South Florida Annenberg Challenge insisted on a dependable and authoritative evaluation of the initiative. They demanded an evaluation that would raise education standards in the community and institutionalize professional evaluation as an integral companion to instructional programs and delivery. This was not the first time the business community had partnered with these school districts. They were not uncommitted, just unconvinced. They refused – and they said this loud and clear to the education community – to blindly invest time, energy, and commitment in the status quo or to become swamped in any bureaucratic quagmire.

Critically important to the success of the evaluation endeavor was the stakeholders' attitude toward the Challenge project. Clearly, the target was to increase student achievement, but they took for granted that there would be hits and misses, midcourse corrections, and redefining of goals. After all, this was an education experiment. If student achievement rose during the process, the strategies could be captured and replicated in other schools. If not, they could be ruled out. What was learned about what works and what doesn't would inform future decisions and action.

Even as SFAC became operational in 1997, Executive Director Elaine Lifton realized that it was time to start taking stock. The initial evaluation process was approved by the SFAC Board in November 1997. The first evaluation team was directed by Jeff Gorrell of Auburn University, with support from a "home" connection at the University of Miami that would conduct on-site visits. Team members were co-investigator Jomills Braddock along with Okhee Lee and Edith Miller from the University of Miami, and Nancy Ares and David Shannon from Auburn.

The first-year evaluation report was issued in December 1998. The team had amassed considerable baseline data about student achievement, school demographics, and school environment. The descriptive data provided a pre-partnership picture of schools within each funded business/school/community partnership, a perspective for interpretation of data collected over the course of the partnerships.

That first report also took note of a shift from the original assumptions in the pro-

BEST COPY AVAILABLE

was used to shape the interview protocols, following Strauss and Corbin's (1990) technique of developing a theory grounded in the data obtained during the study. The interview protocols were open to change over time, which turned out to be an important response to the evolution of the SFAC visions and plans over the years.

Major topics of project implementation were used as the contextual framework to explore the progress of each partnership and the overarching concept of success. Respondents were interviewed about the following topics: project history on the partnership level; project goals; partnership formation and development; project implementation, systemic impact and sustainability; strengths and limitations; and evaluation of the project as it related to administrators, teachers, students, parents, and statewide assessment. Structuring the interviews in this way ensured that most relevant partnership phases were covered, while permitting the respondents a great deal of flexibility in interpreting their experiences.

Interview protocols and case study research were also used to develop survey questionnaires and monitor reports. To this end, a matrix was developed to make certain that all areas under investigation would be probed using multiple data sources. A detailed list of all possible topics, areas for investigation, success indicators, and concepts and issues was created. Common themes were collapsed into larger categories until all questions and probes were classified. When questions or probes were missing from specific instruments or protocols, they were added as necessary. The evaluation team reviewed all the interview questions and probes for relevance, clarity, brevity, and appropriateness (Slavin 1984). We will provide further details about surveys and data monitoring below, in the section Future Directions for Program Evaluation.

posal submitted to the Annenberg Foundation to five focus areas articulated by the board (see text). Part of this evolution was ascribed to the organization's response to the unfolding experiment, but a large part was due to the leadership of the new chair of the SFAC board, Leonard Miller.

Over the next year, the evaluation strategy shifted along with the SFAC vision, and the evaluation team itself also underwent changes. In late 1999, Gorrell left Auburn University, and the SFAC board considered further localizing the evaluation. The Auburn University team members withdrew, and in January 2000 the Board contracted with the University of Miami team (now comprising Ann Bessell, Jeanne Shay Schumm, and Okhee Lee) as the sole evaluators.

The evaluation work now shifted to a greater emphasis on case studies, responding to the board's desire to invest in the most promising partnerships, find out what made them tick, and thus see what to scale up. Recognizing that the partnership initiatives were likely to be a major contribution of the South Florida Annenberg Challenge, they wanted to be sure they understood when it happened well, how it happened, and what could be learned. Being local, the team was able to spend time at the schools and became more involved in identifying and screening promising case study sites. As the case studies yielded timely information about the partnerships, SFAC was able to respond, regroup, and focus funding and effort where it would have the most impact.

The South Florida Annenberg Challenge has generated annual evaluations since 1998 and biannual monitoring reports since 2000. The most recent report, issued in December 2001, not only tracks progress but backs it up with significant evidence that can only be captured over time, including trend lines and longitudinal data for up to three years. The final phase of the evaluation, which will be undertaken in 2003, addresses operational effectiveness of the South Florida Annenberg Challenge effort.

Recalling the journey from fledgling organization to full-fledged, statewide operation, the SFAC board's late chairman, Leonard Miller, wrote:

> The evaluation process and our insistence on data-driven assessment is a hallmark of the South Florida Annenberg Challenge. As information emerged, we were able to focus our programs around these validated lessons and our... mission. The lessons learned – and we stand by them because our evaluation process is solid – will be our legacy.

BEST COPY AVAILABLE

CASE STUDY METHODOLOGY

Case studies were designed to provide an in-depth look at selected SFAC partnerships. The case studies describe project implementation, analyze outcomes, and examine the degree to which the project goals were accomplished. Each case study involved individual and focus group interviews with participants at the partnership, school, family, and student levels, as well as observations of school activities and partnership events. The case studies highlighted perceptions of key stakeholders about project implementation and outcomes. Case study results indicated major factors for effective partnerships, lessons learned from this reform initiative, and recommendations for future research and evaluation. The case studies contained a rich array of data offering insights into ways in which the SFAC had an impact on the tricounty community and the state.

Case study approaches were modified over the years. Initially, case studies were designed to represent five SFAC vision statements espoused by the Executive Board; these statements addressed technology, smaller and more focused schools, rewards and incentives, parent involvement, and systemic change. Later on, case studies were selected with the following criteria: they were representative of one of the three districts and a regional partnership; they were representative of the SFAC vision statements; they had made substantial progress in achieving stated objectives as indicated in monitoring reports; and they had shown promising student-achievement outcomes as demonstrated by statewide achievement-test scores.

The case studies contained a rich array of data offering insights into ways in which the SFAC had an impact on the tricounty community and the state.

Using the five vision statements as the framework, case study research was guided by seven sets of questions: goals of the project (in relation to SFAC vision statements); partnership among business, community, and education; measures of student achievement; systemic impact in terms of sustainability and replicability; budget (in relation to SFAC vision statements); barriers and facilitators; and indicators of success.

In 2000, the SFAC Executive Board changed its five vision statements to four themes leading to student-achievement outcomes: principal leadership, teacher quality, parental involvement, and regional collaboration. Case studies were selected to represent these four themes. Case study research was guided by questions related to stakeholders' understanding of project goals, formation and implementation of partnerships, systemic impact in terms of sustainability and replicability, and barriers and facilitators. In addition, stakeholders' perceptions of principal leadership, teacher quality, and parental involvement were examined.

As the SFAC was implemented over the years, partnerships were at different levels of implementation at any given time. To examine the life span of successful partnerships, case studies were selected at two levels: exploratory investigations of partnerships showing initial evidence of improved student-achievement outcomes and comprehensive investigations of those partnerships that demonstrated improved student-achievement outcomes. Eventually, partnerships with proven records of student achievement will be able to apply for achievement grants, designed to identify and develop strategies to sustain and replicate successful partnerships beyond the span of the SFAC funding.

A total of fourteen case studies were done. The initial four began during 1999; four more began in 2000; and six more began in 2001. This chapter includes the results from thirteen case studies: four from Miami–Dade County, two from Broward County, four from Palm Beach County, and three regional partnerships. (Note: One case study was not continued when the project was terminated due to its lack of progress.)

Demographics

Students' demographic characteristics included the following data for participating schools (reported in percentages): minorities, free/reduced lunch, limited English proficiency (English-language learners), and exceptionalities. SFAC schools' demographic characteristics during the 2000–2001 academic school year were examined in comparison with each school district characteristics (see Table 1).

- In each of the three districts, SFAC schools had higher overall percentages of minority students than the district averages. In individual SFAC schools, the percentage of minority students ranged from 12.5 percent to 100 percent, with an average of 76 percent.
- In each of the three districts, SFAC schools had higher overall percentages of students who were eligible for free and/or reduced lunch than the district averages. In individual SFAC schools, the percentage of eligible students ranged from 5.6 to 99.4 percent, with an average of 61.8 percent.
- In Broward and Palm Beach counties, SFAC schools had higher overall percentages of students who were English-language learners (ELL) than the district averages; in Miami–Dade County, the overall percentage of ELL students in SFAC schools was only 0.3 percent lower than the district average. In individual SFAC schools, the percentage of students who were English-language learners ranged from 0 to 73.7 percent, with an average of 14.8 percent.
- SFAC schools in Miami–Dade and Broward counties had higher overall percentages of students in exceptionalities programs than the district averages; in Palm Beach County, SFAC schools were only 0.7 percent lower than the district average. In individual SFAC schools, the percentage of students who were in exceptionalities programs ranged from 1.4 to 42 percent, with an average of 13.6 percent.

As indicated by the demographic data, the SFAC partnerships benefited student populations that, because of their low socio-economic status and diverse needs, are in greatest need of additional educational support.

		% English-language learners	% Free/Reduced Lunch	% Minorities	% Exceptionalities
Miami-Dade	Overall	18.9	69.0	88.9	10.4
	SFAC	18.6	76.7	91.1	15.4
Broward	Overall	10.1	31.5	61.5	10.9
	SFAC	12.1	51.5	82.5	11.6
Palm Beach	Overall	10.3	32.9	52.9	13.3
	SFAC	11.2	58.1	64.5	12.6

Source: Schumm, Lee & Bassett 2002, p. 3.

Table 1. SFAC schools' demographic characteristics during 2000–2001

Participants

Participants were representative of stakeholders at the partnership and school levels. The exact composition of participants varied among the projects depending on the nature of each partnership. From January 1999 to December 2001, the University of Miami team collected data from 176 individual interviews and sixty-seven focus group interviews. Approximately 580 key stakeholders in the tricounty area participated over a three-year period.

Data Sources and Collection

Data collection efforts were based on the premise that no single data source was likely to be bias-free or to be a completely accurate representation of reality. Researchers used multiple opportunities to explore participants' perspectives among different groups of stakeholders – business partners, community partners, project directors/coordinators, principals, teachers, and parents – and across two levels of investigation. In addition to monitoring reports and grant documents, Level 1 (exploratory) case studies included individual phone interviews with business

and community partners and principals, and focus group interviews with project directors, coordinators, and staff. Level 2 (comprehensive) case studies added focus-group interviews with teachers, parents, and students, and site visits when feasible.

Individual Telephone Interviews

Telephone interviews were conducted with a sample of 176 individuals. Each interview lasted approximately thirty minutes and was audiotaped and transcribed. Researchers' notes were incorporated into the transcribed interviews to ensure a complete and accurate data set. The interview protocol included questions and probes to examine stakeholders' understanding of project goals, formation and implementation of partnerships, systematic impact in terms of sustainability and replicability, and barriers and facilitators. In addition, stakeholders' perceptions of principal empowerment, teacher quality, and parental involvement were examined.

On-Site Focus Group Interviews

There were sixty-seven on-site focus group interviews. Each focus group interview lasted about one and a half hours and followed the same guidelines used for phone interviews. One researcher served as facilitator and a second member of the evaluation team served as note-taker. During many of these visits the principal would provide a tour of the school, focusing on students and teachers who were participating in the project. These on-site visits provided an opportunity to gain additional insight and enhanced the validity and reliability of the subsequent analyses.

As key findings emerged, there was also a continuous search for examples that would contradict the key findings.

Demographics, Achievement-Test Scores, and School Grades

Demographic data included ethnicity, socio-economic status (as reflected by free and/or reduced lunch), school enrollment, average class size, average years of teacher experience, percent of students absent over twenty-one days, stability rate, percent of students qualifying for exceptional student education, percent of students who are English-language learners, incidents of crime and violence, percent of students involved with in-school suspensions, and percent of students in out-of-school suspensions.

Student-achievement data consisted of scores on the Florida Comprehensive Assessment Test (FCAT) for reading, writing, and mathematics. School grades were based on the Governor's A+ Plan for Education developed by the Florida Department of Education. School grades are determined by several criteria established by the department including FCAT achievement, absenteeism, and drop-out rates. School grades were used to identify high- and low-performing schools, stimulate academic improvement, and summarize information about school achievement. Students' FCAT results and school grades are beyond the scope of this chapter and are not reported here. The results are available in the 2001 SFAC evaluation report (Schumm, Lee & Bessell 2002).

Data Analysis

Each case study was assigned a team of two researchers to ensure consistency in data collection methods, opportunities for ongoing exchange of emerging ideas, and inter rater reliability. In addition, reliability of the qualitative portion of this evaluation was addressed using standard procedures in the field, such as defining and reporting methods, using a framework to guide the evaluation, and using multiple evaluators (Seale 1999).

All transcripts from individual phone interviews and focus group interviews were coded and analyzed using the qualitative analysis software ATLAS/ti. A coding system was developed and themes were identified under each of the five major categories, including collective participation, principal leadership, teacher quality, flexible and innovative curriculum and finance, and parental involvement.

A two-step process was used for data analysis. First, for each category, a theme was identified and specific evidence was noted (e.g., quotes and location in transcript). Then, overall patterns were identified across interviews and participants. After the analysis of each case study was completed, the results of the case studies were examined for commonalities and differences among the projects. As key findings emerged, there was also a continuous search for examples that would contradict the key findings (Yin 1994). To increase reliability, external checks were conducted using a second researcher who independently practiced coding the same transcripts using an analysis framework. When an inconsistency arose, the researchers discussed and negotiated until they reached consensus. When reliability between coders reached 90 percent, the coders continued analysis independently.

Based on cross-case analysis, important factors leading to a successful project were identified. These factors were framed in terms of a "profile" rather than a "definition," since there were great variations in project goals, objectives, and level of implementation. It is to be noted that these results were based largely on the perceptions of participants who were representative of business, community, and education partners.

HIGHLIGHTS OF CASE STUDY RESULTS

With few exceptions, individual partnerships made great strides in building school/business/community collaboration. In many cases, partnerships represented first-time efforts in collaboration among these sectors. Although key stakeholders worked largely in uncharted waters, they managed to discover how best to communicate with one another, launch their projects, and keep their projects afloat. Stakeholders in leadership positions recognized the pivotal role of the SFAC in forming these partnerships.

The evidence from case studies indicated key themes about successful partnerships that fell into the five major analysis categories. Key characteristics of each category are described here, and a partnership to illustrate these characteristics is presented.

Although key stakeholders worked largely in uncharted waters, they managed to discover how best to communicate with one another, launch their projects, and keep their projects afloat.

Collective Participation

The SFAC was a catalyst in forming business, community, and education partnerships. Case studies reflected a high level of satisfaction with SFAC projects among key stakeholders. The majority agreed that their project was a good example of collective participation and should be continued in their school and replicated in other schools as well. The following characteristics of collective participation provided an atmosphere for success.

- Participants shared clear, common goals, which led to focused vision for the school's reform efforts.
- Common language developed, facilitating mutual respect, communication, and constructive feedback among participants. This challenge often included dropping the use of acronyms common in either academic environments or business interactions.
- The partnership was actively involved and visible in the school and community. This not only provided community awareness but also added to a positive environment, mutual buy-in, and commitment to the project.
- Participants were willing to relinquish individual autonomy and to collaborate and use diversity of talents among participants. It was important to leave egos at the door and focus on what would be best for the students.

Partnerships with consistent leadership of the same principals exhibited increased productivity and participant buy-in.

- Participants, particularly principals and teachers, felt ownership of the project and felt empowered to facilitate implementation.
- Partners displayed sensitivity and realistic understanding of time commitments. Principals and teachers often had multiple responsibilities requiring time, paperwork, and accountability.
- Strong leadership at each level of the state, district, partnership (i.e., project director), and school (i.e., principals and lead teachers).

Central EXPRESS

This partnership attempted to improve student achievement in a systemic way. The focus was to provide comprehensive and consistent teacher professional development and to develop a "seamless" educational experience for children from elementary school through high school within a single feeder pattern. Teachers began their collaboration across schools in the area of writing and saw marked improvements in standardized-writing-test scores in all participating schools. Plans to work on reading achievement using similar models of teacher professional development and communication with key stakeholders are in progress. Leadership in Central EXPRESS was successful in garnering resources from local and national businesses as well as local institutions of higher education.

Principal Leadership

Principals played key roles in successful projects. Partnerships with consistent leadership of the same principals exhibited increased productivity and participant buy-in. This was enhanced when the principals actively participated in training, were available and committed to resolve barriers as they arose, and maintained visibility for both staff and students during project implementation. Effective principals displayed the following characteristics:

- stability (low turnover): Multiple leaders over the life of a project can lead to fragmented leadership and lack of commitment or buy-in
- focus on academic and instructional goals to improve student achievement
- active involvement and visibility of principles at the school and in the community
- attention to local norms and concerns
- promoting collaboration for mutual benefit with/among teachers
- shared decision making with teachers
- support for buy-in of teachers

Partnership to Advance School Success (PASS)

PASS's success was founded on a strong school/community/business partnership and total commitment of key stakeholders that emphasized principal leadership. A principal, an educational coach, and a corporate CEO formed a team for each school and worked as partners to create and implement individual reform initiatives tailored to meet the school's unique needs for improvement with the goal of ultimately raising the school grade to an "A" status. Schools focused on such areas as specific student-achievement goals, streamlined planning, and budget structuring. The consistent support and assistance of the principal-coach-CEO triads were integral to the progress of the schools toward meeting their established goals.

PASS schools showed a trend of steady increases in student achievement. From 1998 to 2001, the four pilot schools showed consistent improvement in standardized-test scores across all three subject areas of reading, writing, and mathematics. The

PASS partnerships provided a valuable platform for school and business professionals to share innovative ideas, alter perceptions of public education, and facilitate school reform. Even before the project was extended outside of the four pilot schools, key stakeholders expressed a strong desire to continue the project and expand it to include more grade levels, other subject areas, and additional schools. Eventually there will be as many as thirty schools throughout the state participating in PASS partnerships.

Teacher Quality

Professional development activities were critically important to promote teacher quality. Effective professional development activities provided opportunities for teachers to implement creative and innovative curriculum materials and instructional strategies, which eventually led to increased student achievement as measured by standardized tests. Effective professional development activities were indicated by the following characteristics:

- focused and relevant to student outcomes, particularly those areas tested on standardized tests
- opportunities for teachers' self-reflection
- opportunities for collaboration and sharing among teachers
- flexibility and support for innovations and creative approaches to teaching

Broward Academy for Teacher Excellence (BATE)
The teacher shortage is a nationwide challenge. BATE represented a broad attempt to combat this shortage on two fronts: by retaining teachers in the profession and by attracting more teachers to the profession. BATE accomplished this by according greater respect to the career through professional development activities. BATE offered an opportunity for schools to serve as professional development demonstration classrooms using action research to improve student achievement. Students in BATE schools benefited from improved and creative instruction. In fact, all four of the schools in the initial phase of the project had standardized-test scores that were higher than the district mean in reading and math.

Flexible and Innovative Curriculum and Finance

The SFAC promoted innovative and creative curriculum materials and instructional approaches to enhance student learning. The SFAC also promoted flexibility in finance and the education system, such as funding opportunities for teacher professional development, offering incentives and rewards for excellence, providing alternative ways to overcome bureaucratic barriers in the education system, and encouraging the establishment of new policies in the district and the school. Flexible, innovative, and unique characteristics of SFAC included

- innovative and creative curriculum and instruction
- flexible funding and management
- support for new ways of thinking "outside the box"
- strategies for coping with bureaucracy and regulations in the education system that sometimes hinder education innovations
- prestige of the Annenberg reputation

Math Is Not Difficult (MIND)
Project MIND provided consistent resources for administrators and teachers to revise mathematics instruction and to develop innovative and creative math-based activities. Students were encouraged to develop original work, such as math songs, poems, art, stories, games, puzzles, and brainteasers, related to math concepts. There was so much excitement about the project that nonparticipating community members heard about

The consistent support and assistance of the principal-coach-CEO triads were integral to the progress of the schools toward meeting their established goals.

Project MIND and wanted to get involved with the strategies. Professional development for teachers and parent involvement were critical components of the program. Other partnerships learned from this innovative initiative that benefited in terms of school/business partnerships, curricular change, and home/school relationships.

Urban Institute for Environmental Studies.
This partnership provided an innovative design for learning to meet student and community needs. The program immersed nonmagnet students in grades K–12 in a learning community that emphasized personalization, intellectual focus, and work-based career learning, resulting in significant improvement in student achievement. Standardized-test scores in reading, math, and writing showed steady improvement over three years. Students attended school more often and had fewer discipline referrals than before the project began, and the high school drop-out rate was reduced. The benefits of specially designed curricula and work-based career learning for low-performing students are additional elements of this highly successful, innovative approach to reaching at-risk students.

Parental Involvement

Compared to other stakeholders, parents were generally less actively involved in SFAC projects. Even in successful projects, parents often indicated that they were not aware of business and community partners or specific project activities. Most projects needed to provide opportunities for parents' participation and awareness. Parental involvement increased in the following situations:

- involvement in children's schoolwork at home
- involvement to promote student outcomes academically and socially
- frequent communication and opportunities to be involved in project/school activities
- user-friendly approaches to meet parents' needs (e.g., evening events, child care)

Family Tech
This partnership represented an ambitious effort to provide equitable access to technology for the most needy children and parental involvement in their children's education. All students in each targeted classroom received computers to take home along with training for students, their families, and their teachers, so that everyone could learn technological skills. Opportunities for computer technology were also incorporated across the curriculum. This project represented an effort to provide professional development for teachers in the critical area of educational technology and to enhance parental involvement by providing parents with opportunities to improve or acquire computer skills while working with their children on school tasks.

Students attended school more often and had fewer discipline referrals than before the project began, and the high school drop-out rate was reduced.

KEY LESSONS LEARNED

Case studies implemented in the formative years of the SFAC helped to shed light on key aspects of this tricounty school reform initiative. The case studies also paved the way for subsequent program evaluation activities. This section provides key recommendations from the initial case studies and concludes with an overview of ongoing and future evaluation effort.

Recommendations for Business/Community/Education Partnerships

The initial case studies indicated that stakeholders viewed the SFAC as a catalyst in forming business, community, and education partnerships. Establishment of this collective participation was an accomplishment in and of itself. Certain characteristics of the SFAC also made it unique in its contribution to school

reform: support for innovative education programs, incentives and rewards for excellence, flexibility with the SFAC funding, and the prestige of the Annenberg reputation. Key stakeholders agreed that their projects would not have been possible without the support of SFAC. This was particularly true in schools with large numbers of minority students and children of poverty that had limited resources for new initiatives.

Thus, the SFAC made unique contributions for business, community, and education sectors to come together and form partnerships for school reform. Stakeholders, including principals, teachers, parents, and students, expressed a high level of satisfaction with their partnerships. The majority agreed that their project was a good example of collective participation and should be continued in their school and replicated in other schools as well. However, most participants did not have specific plans or funding sources to sustain their projects. Strategies to sustain and replicate effective partnerships require consideration.

Principals were essential for program implementation and success. Principals served two major roles: to translate the vision and goals of the partnership into practice and to be a primary communication link among various groups of participants. Principal buy-in was one of the most important factors for program success. If the principal did not buy into the project, the project floundered at that school, which led to wide variations among schools within the same partnership. In addition, the project director had a tremendous hurdle to cross without the principal's support.

Teacher buy-in was critical for successful implementation of project activities at the classroom level. Shared decision-making practices empowered teachers and increased their feelings of ownership and investment in the project. In addition, effective professional development activities provided opportunities for teachers to implement creative and innovative curriculum materials and instructional strategies. Initial findings indicated that traditional workshops and limited professional development activities did not provide teachers with the tools they needed to teach students in urban settings. Effective approaches for teacher professional development for more intensive instruction in urban settings are needed.

Principal buy-in was one of the most important factors for program success.

A key characteristic of the SFAC was its support for innovative, creative, and flexible education programs. The SFAC also promoted flexibility in finance, such as funding opportunities for teacher professional development and incentives and rewards for excellence. In addition, the SFAC was a catalyst for partnerships to develop strategies for overcoming bureaucracy and regulations in the education system that sometimes hindered education innovations. Unfortunately, procedures for deregulation at the state and district levels were not consistent, making successful program implementation difficult. The education system and some educators seemed to resist changes in the system.

Parents were key partners in public education. This support was especially important in inner-city schools serving minority students from low socioeconomic backgrounds. Successful partnerships promoted parental involvement in academically oriented activities. Compared to other stakeholders, however, parents were generally less actively involved in SFAC projects. Even in successful projects where parental involvement was considered a priority, parents often indicated that they were not aware of business and community partners or specific project activities.

The results indicate that to launch a complex partnership, leadership of key personnel is critical. At the partnership level, the participation of the project director was important. At the school level, principal support was necessary. Commitment of key selected participants, such as lead teachers, project

It was necessary to give up individual autonomy and be willing to collaborate and use the diversity of talents among participants.

coordinators, or technology facilitators, was necessary, particularly at initial stages of program implementation. In addition, a sense of ownership and empowerment of participants, particularly principals and teachers, was critically important. The various groups of people needed to work together toward achieving a clear set of common goals. As mentioned above, it was necessary to give up individual autonomy and be willing to collaborate and use the diversity of talents among participants. It was also important to be sensitive and realistic about time commitments.

Future Directions for Program Evaluation

The initial case studies contain a rich array of data offering insights into the ways in which the SFAC had an impact on the community. Current SFAC case study research comprehensively involves multiple qualitative and quantitative data sources. Individual and focus-group interviews continue to serve as the core of case studies. Recently, the evaluation team developed a rating rubric to summarize the interview data in a quantitative manner. In addition, survey results with larger numbers of stakeholders, monitoring reports, more refined student-achievement-test analyses, and cost-effectiveness analyses are included as part of the case study reports.

Rating Rubric

To further refine the analysis of individual and focus-group interviews, a rating rubric was developed. Key indicators of success, gleaned from relevant literature and our previous case studies, were combined with survey results, monitoring reports, and lessons learned to create the units of analysis for a rating rubric. The rubric included salient components for the following seven constructs: collective participation, principal leadership, professional development/skilled teachers, parent involvement, student outcomes, SFAC project staff, and school environment. These constructs were further defined in subcategories. As a result, each transcript could be rated on a four-point scale, using the rating rubric, to determine a score for each construct. Mean scores for each category could be determined, allowing comparisons across all partnerships and across schools within each partnership of the seven constructs separately and collectively.

Surveys

As case study research continued, it became necessary to examine whether the case study results with selected partnerships were replicable with all partnerships. Surveys were conducted with key stakeholders from all partnerships in spring 2001 and 2002. Construction of survey questionnaires was based on two primary data sources: indicators of success developed from the cross-case analysis of case studies, and indicators of success gleaned from an extensive review of the literature. The evaluation team conducted an extensive review of literature on key indicators of principal leadership, teacher quality, and parental involvement that lead to student achievement. Key indicators of success from the literature were combined with key indicators of success from the ongoing case studies of partnerships.

The results from these two sources were generally consistent, although the case study results provided additional insights into business/community/school partnerships in urban school settings. Survey subscales were created for the same seven constructs of success used to evaluate interview transcripts, enabling us to look at overall means from both surveys and case study partnership interviews and, in addition, providing a way to determine which constructs contribute most to success at both the school and partnership levels.

Eventually, interview protocols for case studies of selected partnerships and survey questionnaires for all partnerships were developed using the same sets of key indicators about collective partnership, principal leadership, teacher quality, and parental involvement. These indicators of success were also used for the biannual monitoring reports described in the next section.

Monitoring Reports

Monitoring reports were designed to provide biannual updates to the SFAC regarding project activities and progress toward achieving goals and objectives, as well as issues or obstacles encountered during project implementation. The reports, usually completed by the project directors, provided a different perspective on the implementation and outcomes of SFAC partnerships.

Each monitoring report included a demographics section, in which partnerships described the schools and partners involved, the number and grade levels of the students, and the project focus (e.g., mathematics, reading, writing, professional development, FCAT, technology, parental involvement, and principal leadership). Next, a free-response section asked the partnerships to respond to questions about key accomplishments, challenges, modifications of project activities, progress in reaching objectives, and determination of the number of key success indicators met. Finally, a short survey section was included to obtain information about the project directors' perceptions of the presence or absence of key indicators of success within the partnership.

Monitoring reports provided an opportunity for each project to communicate successes and lessons learned and provide unique insights into the dynamics of the school, business, and community partnerships. In addition, these reports served as important links between each project and the SFAC office. The review and analysis of the biannual monitoring reports allowed the SFAC staff to become aware of areas in which each project might need assistance or input and ensured that each partnership was making the expected progress.

Student Achievement

Student-achievement data (FCAT scores in reading, mathematics, and writing) were analyzed for all SFAC schools and were included in case study reports. Comparison schools (in terms of demographics and initial FCAT scores at baseline) for all SFAC schools have been identified. The FCAT scores for SFAC partnerships and schools within each partnership were analyzed and compared to those of the comparison schools, the respective school district, and the state overall.

Cost-Effectiveness Analysis

To assist the SFAC in identifying partnerships that not only appear to be effective in improving student outcomes, but that also provide a reasonable return on investment, a cost-effectiveness analysis plan is in development. This initial attempt is a systematic "broad-stroke" approach. The results will provide practical guidance for identifying partnerships that are both academically successful and financially efficient.

Levin and McEwan (2001) expressed concern about the quantity and quality of cost-effectiveness studies. It is often difficult to determine accurate costs and to develop good measures of cost-effectiveness. A set of assumptions needs to be made when determining cost per student and calculating cost-effectiveness, since it is not feasible to conduct an on-site audit at each school. The resulting cost-effectiveness ratio can be expressed in two ways: to

Monitoring reports were designed to provide biannual updates to the SFAC regarding project activities and progress toward achieving goals and objectives, as well as issues or obstacles encountered during project implementation.

determine which projects spent the least amount of money per student (cost) for a given level of effectiveness (gain in standardized-test scores) and to examine the highest level of effectiveness for a given cost per student (Levin & McEwan 2002).

A FINAL WORD

The SFAC took the approach of encouraging grassroots reform initiatives to scatter seeds and see what grows, rather than targeting a specific goal (e.g., smaller class size) as the driving theme. Such an approach may be appropriate in school districts that have a longer history of school reform efforts or philanthropic funding for private-public partnerships; school reform, particularly at a multidistrict level, is new to South Florida. Further research is needed to assess the pros and cons of this approach and to document lessons learned in order to guide other large-scale school reform initiatives.

The findings from the case studies suggested that factors associated with principal leadership, teacher empowerment, and parental involvement in private-public partnerships can lead to increased student achievement. The results also indicated barriers to success. The study of the South Florida Annenberg Challenge offers important insights and valuable lessons to the public, as well as educators, about the civic-minded approach to school reform. Others interested in forming similar partnerships or achieving similar goals may be able to learn from difficulties that participants in these partnerships faced and the solutions that they discovered.

References

Bracey, G. W. 1999. "The Ninth Bracey Report on the Condition of Public Education," *Phi Delta Kappan* 81:2, 147–68.

Breen, R., A. Jenkins, R. Lindsay, and P. Smith. 1998. "Insights through Triangulation: Combining Research Methods to Enhance the Evaluation of IT Based Learning Methods." In *Innovations in the Evaluation of Learning Technology*, edited by M. Oliver. London: University of North London Press.

Fitz, J., and S. Gorard. 2000. "School Choice and SES Stratification of Schools: New Findings from England and Wales." Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Fullan, M. 2000. "The Three Stories of Education Reform," *Phi Delta Kappan* 81:8, 581–584.

Gay, L. R., and P. Airasian. 2000. *Educational Research: Competencies for Analysis and Application*. 6th ed. Upper Saddle River, NJ: Prentice-Hall.

Hoyle, J. R., and R. O. Slater. 2001. "Love, Happiness, and America's Schools: The Role of Educational Leadership in the Twenty-First Century," *Phi Delta Kappan* 82:10, 790–795.

Levin, H. M., and P. J. McEwan. 2001. *Cost-Effectiveness Analysis: Methods and Applications*. 2nd ed. Thousand Oaks, CA: Sage.

Levin, H. M., and P. J. McEwan. 2002. *Cost-Effectiveness and Educational Policy*. Larchmont, NY: Eye On Education.

McDonald, J. P. 1999. "The Trouble with Policy-Minded School Reform and Why It Should Give Way to a Civic-Minded Alternative," *Education Week* (September 8): 68, 44, 47.

Morgan, D. L., and R. A. Krueger. 1993. "When to Use Focus Groups and Why." In *Successful Focus Groups: Advancing the State of the Art*, edited by D. L. Morgan. Newbury Park, CA: Sage.

Posner, G. J. 1995. *Analyzing the Curriculum*. 2nd ed. New York: McGraw-Hill.

Rallis, S. F., and M. M. MacMullen. 2000. "Inquiry-Minded Schools: Opening Doors for Accountability," *Phi Delta Kappan* 81:10, 766–773.

Rubin, H. J., and I. S. Rubin. 1995. *Qualitative Interviewing: The Art of Hearing Data*. Thousand Oaks, CA: Sage.

Schumm, Jeanne Shay, Okhee Lee, and Ann G. Bessell. 2002. *South Florida Annenberg Challenge Evaluation Report: January 1, 2001 to December 31, 2001*. Coral Gables, FL: University of Miami Press.

Seale, C. 1999. *The Quality of Qualitative Research*. Thousand Oaks, CA: Sage.

Slavin, R. E. 1984. *Research Methods in Education: A Practical Guide*. Englewood Cliffs, NJ: Prentice-Hall.

Strauss, A., and J. Corbin. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage.

Tashakkori, A., and C. Teddlie. 1998. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage.

Yin, R. K. 1994. *Case Study Research Design and Methods*. Thousand Oaks, CA: Sage.

APPENDIX

Cross-Site Research Reports on the Annenberg Challenge

A Challenge to the Nation: A Progress Report
October 1995

A Challenge to the Nation: A Progress Report
December 1996

A Challenge to the Nation: A Progress Report
February 1998

Doing What You Mean to Do in School Reform: Theory of Action in the Annenberg Challenge
(Annenberg Institute Occasional Paper no. 4)
Donald A. Schön and Joseph P. McDonald
June 1998

Citizens Changing Their Schools: A Midterm Report of the Annenberg Challenge
April 1999

Meeting the Challenge: Fundraising Progress in the Annenberg Challenge
January 2000

Students at Work: A Portfolio from the Annenberg Challenge
April 2000

The Annenberg Challenge: Lessons and Reflections on Public School Reform
June 2002

The Arts and School Reform: Lessons and Possibilities from the Annenberg Challenge Arts Projects
February 2003

Research Perspectives on School Reform: Lessons from the Annenberg Challenge
March 2003

Final Evaluation Reports on the Annenberg Challenge Sites

Atlanta

Urban Atlanta Coalition Compact Annual Evaluation Report, 2000–2001
July 2001
Georgia State University Evaluation Team
Information at Georgia State University, Department of Psychology

Bay Area

Bay Area School Reform Collaborative: Phase One (1996–2001) Evaluation
August 2002
Milbrey W. McLaughlin and Joan E. Talbert
Center for the Context of Teaching at Stanford University
Ordering information at *www.basrc.org*

133

Boston

Taking Stock: The Status of Implementation and the Need for Further Support in the BPE–BAC Cohort I and II Schools
May 2001
Barbara Neufeld and Katrina Woodworth
Education Matters, Inc.
Available at *www.bpe.org/searchpubs.aspx*

Chattanooga

Creating a New, Standards-Based School System: Final Report of the Hamilton County Annenberg Challenge, 1995–2001
September 2001
Patricia Ceperley and Pamelia Coe
AEL, Inc.
Information at *www.pefchattanooga.org*

Chicago

Improving Chicago's Schools, 1996–2001: Effective Strategies for Change
Forthcoming
Mark A. Smylie, Stacy A. Wenzel et al.
Consortium on Chicago School Research
Available at *www.consortium-chicago.org*

Detroit

Schools of the 21st Century –
Detroit Annenberg Challenge
Final Evaluation Report forthcoming (2003)
SAMPI–Western Michigan University, Abt Associates, Roegan Enterprises
Information at *www.s21c-detroit.org*

Houston

Houston Annenberg Challenge
Final Evaluation Report forthcoming (2003)
Pedro Reyes and Joy C. Phillips
University of Texas at Austin with University of Houston and Rice University
Information at
www.utexas.edu/projects/annenberg/index.html

Los Angeles

The Los Angeles Annenberg Metropolitan Project: Evaluation Findings
March 2003
Joan L. Herman and Eva L. Baker
National Center for Research on Evaluation, Standards, and Student Testing at the University of California at Los Angeles
Available at *http://CRESST.org*

Minneapolis

Arts for Academic Achievement
Final Evaluation Report forthcoming (March 2003)
Center for Applied Research and Educational Improvement at the University of Minnesota
Information at *http://education.umn.edu/CAREI*

New York Arts

Final Evaluation Report on The Center for Arts Education's New York City Partnerships for Arts and Education Program
November 2001
Terry Baker, Bronwyn Bevan, and Noga Admon
Center for Children and Technology at Education Development Center
Available at *www.cae-nyc.org/Downloads/cae2002.pdf*

New York Networks

Final Report of the Evaluation of New York Networks for School Renewal: An Annenberg Foundation Challenge for New York City, 1996–2001
December 2001
Dana Lockwood, Norm Fruchter et al.
Institute for Education and Social Policy at New York University
Available at *www.nyu.edu/iesp/publications/reports.html*

Philadelphia

The Limits and Contradictions of Systemic Reform: The Philadelphia Story
November 2002
Tom Corcoran and Jolley Bruce Christman
Consortium for Policy Research in Education at the University of Pennsylvania
Available at *www.cpre.org/Publications*

Rural

Living and Learning in Rural Schools and Communities / Lessons from the Field (2 vols.)
February 1999
Annenberg Rural Challenge Research and Evaluation Program
Information at *www.ruraledu.org*

Salt Lake City

Eccles/Annenberg Challenge
Final Evaluation Report forthcoming (March 2003)
University of Utah
Information at *www.slc.k12.ut.us*

South Florida

South Florida Annenberg Challenge
Final Evaluation Report forthcoming (January 2004)
University of Miami School of Education
Information at *www.changeeducation.org*

TETAC

Final Evaluation Report of the Transforming Education Through the Arts Challenge
July 2002
Joy Frechtling, Sandra Rieder, Joan Michie, Kyle Snow, Philip Fletcher, Jean Yan, and Atsushi Miyaoka
Westat
Information at *www.arts.ohio-state.edu/naec*

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Research Perspectives on School Reform: Lessons from the Annenberg Challenge	
Author(s): Brenda Turnbull, Editor	
Corporate Source: Annenberg Institute for School Reform at Brown University	Publication Date: March 2003

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

[X]

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

[ ]

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

[ ]

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Susan C. Fisher	Printed Name/Position/Title: Susan Fisher, Associate Director, Comm.
Organization/Address: Box 1985 Brown University Providence, RI 02912	Telephone: 401-863-3552 FAX: 401-863-1290
	E-Mail Address: Susan_Fisher@Brown.edu Date: 11/24/03

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, pleas provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is public available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly mor stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name an address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Urban Education
Box 40, Teachers College, Columbia University
New York, NY 10027

Telephone: 212-678-3433
Toll Free: 800-601-4868
Fax: 212-678-4012

WWW: http://eric-web.tc.columbia.edu

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document bein contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)